



UNIVERSIDAD DE MÁLAGA
FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

**PREDICCIÓN DE INSOLVENCIA EN LOS
SECTORES ECONÓMICOS: UN ANÁLISIS
COMPARATIVO**

Doctorando: Gonzalo Laguillo Díaz

Tesis Doctoral


Dirigida por:
Dr. Manuel Ángel Fernández Gámez
Dr. Rafael Becerra Vicario

NOVIEMBRE DE 2015



Publicaciones y
Divulgación Científica

AUTOR: Gonzalo Laguillo Díaz

 <http://orcid.org/0000-0002-8447-1322>

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional:

Cualquier parte de esta obra se puede reproducir sin autorización pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer obras derivadas.

<http://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

Esta Tesis Doctoral está depositada en el Repositorio Institucional de la Universidad de Málaga (RIUMA): riuma.uma.es

MANUEL ÁNGEL FERNÁNDEZ GÁMEZ y RAFAEL BECERRA VICARIO, ambos Profesores Titulares del Departamento de Finanzas y Contabilidad de la Universidad de Málaga.

Certificamos:

Que bajo nuestra dirección, D. Gonzalo Laguillo Díaz, licenciado en Ciencias Económicas y Empresariales (Sección: Económicas), ha realizado el trabajo de investigación correspondiente a su tesis doctoral titulada:

**“PREDICCIÓN DE INSOLVENCIA EN LOS SECTORES
ECONÓMICOS: UN ANÁLISIS COMPARATIVO”**

Revisado el mismo, estimamos que puede ser presentado al Tribunal que ha de juzgarlo.

Y, para que conste a efectos de lo establecido en la Normativa vigente, autorizamos la presentación de esta Tesis en la Universidad de Málaga.

Málaga a 10 de noviembre de 2015.

Fdo. Manuel Ángel Fernández Gámez

Fdo. Rafael Becerra Vicario

*"Si cada día coges un puñado de arena,
formarás una montaña"*

Buda Gautama

Quisiera comenzar dando las gracias a mis directores de tesis, Dr. D. Manuel Ángel Fernández Gámez y Dr. D. Rafael Becerra Vicario, por su orientación y dedicación en todo momento a lo largo de este trabajo.

Particularmente al profesor Manuel Ángel Fernández Gámez por su manera de trabajar conmigo. Su ánimo, paciencia y exigencia han sido esenciales para el desarrollo de este trabajo y para mi formación como investigador.

Quiero dedicar este trabajo especialmente a mi familia. A mi mujer, Marina, por su amor y su total apoyo durante todos estos años en todo lo que he querido emprender. A nuestras hijas, Lucía e Inés, que son nuestra ilusión y nuestra alegría, a las que me gustaría transmitir, a través de este trabajo y las horas dedicadas al mismo, el valor que tiene la perseverancia y el esfuerzo. Ellas lo son todo para mí.

A mis padres, Antonio y Josefina que, aunque se marchó de nuestro lado en 2004, sigue viva en nuestro recuerdo. Ellos, entre otras muchas cosas, me inculcaron la importancia de la educación, la constancia y la formación. Honrarles con este trabajo ha sido un motor muy importante para mí y una pequeña muestra de mi gratitud como hijo. Igualmente quiero tener aquí un recuerdo para mi hermano Ignacio y su familia, con la convicción y la esperanza de que, al final, después de la tormenta, siempre escampa.

Tampoco puedo olvidar a mis suegros, Antonio e Inés, ni a mis cuñadas, Esther y Celia, que me acogieron como un hijo y un hermano más.

Mi querido amigo Antonio Ayala siempre me animó en este proyecto y me dio valiosos consejos.

Por último, deseo mostrar mi gratitud a mis compañeros doctorandos del departamento de Finanzas y Contabilidad por el apoyo que me han dado. Muy especialmente a David Alaminos, por su colaboración en el trabajo empírico, sin olvidar a otras personas que, de una manera u otra, me ayudaron a aclarar conceptos o me facilitaron su asistencia en alguna parte de este trabajo. En concreto, quisiera mostrar mi agradecimiento a Joaquín Muñoz García y Rafael Ayala Gómez, ambos profesores de Estadística y de Matemáticas, respectivamente, de la Universidad de Sevilla.

En definitiva, gracias a todas aquellas personas que me impulsaron y animaron para que pudiera culminar este proyecto tan deseado por mí.

INDICE

	Pág.
INTRODUCCIÓN	23
 <u>PARTE I: INVESTIGACIÓN DOCUMENTAL Y METODOLÓGICA</u>	31
 CAPÍTULO 1. LA INVESTIGACIÓN PREVIA EN PREDICCIÓN DE INSOLVENCIA.	33
1.1 INTRODUCCIÓN	35
1.2 LA PREDICCIÓN DE INSOLVENCIA EN LAS EMPRESAS ECONÓMICAS	36
1.3 MODELOS GLOBALES DE PREDICCIÓN DE INSOLVENCIA	49
1.4 MODELOS DESCENTRADOS Y MODELOS CENTRADOS	52
1.4.1 MODELOS DESCENTRADOS	55
1.4.2 MODELOS CENTRADOS	98
1.4.2.1 Sector Agricultura	98
1.4.2.2 Sector Industria	104
1.4.2.3 Sector Construcción	118
1.4.2.4 Sector Comercio y Servicios	127
1.4.2.5 Sector Hostelería	129
1.4.3 ALGUNAS CONCLUSIONES ACERCA DEL DESARROLLO DE LA LITERATURA SOBRE PREDICCIÓN DE INSOLVENCIA	137
1.5 FORMULACIÓN DE HIPÓTESIS	157
 BIBLIOGRAFÍA CAPÍTULO 1	161

	Pág.
CAPITULO 2. MÉTODOS DE CLASIFICACIÓN BINARIOS	199
2.1. ANÁLISIS DISCRIMINANTE	202
2.2. MODELOS DE RESPUESTA CUALITATIVA	208
2.2.1 MODELO LINEAL DE PROBABILIDAD	211
2.2.2 MODELOS DE PROBABILIDAD NO LINEAL	213
2.2.2.1 Regresión Logística	214
2.2.2.2 Criterios para la selección de modelos estadísticos	218
2.2.2.2.1 Método de Validación Cruzada (CV)	219
2.2.2.2.2 Criterios basados en la información estadística	220
2.2.2.2.2.1 Criterio de Akaike (AIC)	221
2.2.2.2.2.2 Criterio de Schwarz (BIC)	223
2.2.2.2.2.3 Criterio de Hannan-Quinn (HQC)	225
 BIBLIOGRAFÍA CAPÍTULO 2	 227
 <u>PARTE II: INVESTIGACIÓN EMPÍRICA</u>	 233
 CAPITULO 3. DATOS Y VARIABLES	 235
3.1 DATOS	238
3.1.1 OBTENCIÓN DE LA BASE DE DATOS	239
3.1.2 OBTENCIÓN DE LAS MUESTRAS	245

	Pág.
3.2 VARIABLES	248
3.2.1 PROCEDIMIENTO DE SELECCIÓN DE LAS VARIABLES	248
3.2.2 VARIABLES EMPLEADAS EN LOS MODELOS DESCENTRADOS	249
3.2.3 VARIABLES EMPLEADAS EN LOS MODELOS CENTRADOS	250
 BIBLIOGRAFÍA CAPÍTULO 3	 257
 CAPITULO 4. RESULTADOS	 261
4.1. INTRODUCCIÓN	263
4.2 RESULTADOS CON INFORMACIÓN DE UN AÑO ANTES DE LA QUIEBRA (M.I)	265
4.2.1 RESULTADOS MODELO GLOBAL	265
4.2.1.1 Análisis Exploratorio	265
4.2.1.1.1 Análisis Descriptivo	265
4.2.1.1.2 Análisis de Normalidad	268
4.2.1.1.3 Tests no paramétricos	270
4.2.1.1.4 Análisis de correlación	272
4.2.1.2 Análisis Confirmatorio	274
4.2.1.2.1 Modelo Logit Global sin variables cualitativas	276
4.2.2 RESULTADOS MODELO AGRICULTURA	283
4.2.2.1 Análisis Exploratorio	283
4.2.2.1.1 Análisis Descriptivo	283
4.2.2.1.2 Análisis de Normalidad	285

	Pág.
4.2.2.1.3 Tests no paramétricos	286
4.2.2.1.4 Análisis de correlación	287
4.2.2.2 Análisis Confirmatorio	289
4.2.2.2.1 Modelo Logit sector agricultura	289
4.2.3 RESULTADOS MODELO INDUSTRIA	293
4.2.3.1 Análisis Exploratorio	293
4.2.3.1.1 Análisis Descriptivo	293
4.2.3.1.2 Análisis de Normalidad	295
4.2.3.1.3 Tests no paramétricos	296
4.2.3.1.4 Análisis de correlación	297
4.2.3.2 Análisis Confirmatorio	300
4.2.3.2.1 Modelo Logit sector industria	300
4.2.4 RESULTADOS MODELO CONSTRUCCIÓN	303
4.2.4.1 Análisis Exploratorio	303
4.2.4.1.1 Análisis Descriptivo	303
4.2.4.1.2 Análisis de Normalidad	305
4.2.4.1.3 Tests no paramétricos	306
4.2.4.1.4 Análisis de correlación	307
4.2.4.2 Análisis Confirmatorio	309
4.2.4.2.1 Modelo Logit sector construcción	309
4.2.5 RESULTADOS MODELO COMERCIO Y SERVICIOS	312
4.2.5.1 Análisis Exploratorio	312

	Pág.
4.2.5.1.1 Análisis Descriptivo	313
4.2.5.1.2 Análisis de Normalidad	314
4.2.5.1.3 Tests no paramétricos	315
4.2.5.1.4 Análisis de correlación	316
4.2.5.2 Análisis Confirmatorio	318
4.2.5.2.1 Modelo Logit sector comercio y servicios	318
4.2.6 RESULTADOS MODELO HOSTELERÍA	321
4.2.6.1 Análisis Exploratorio	321
4.2.6.1.1 Análisis Descriptivo	321
4.2.6.1.2 Análisis de Normalidad	322
4.2.6.1.3 Tests no paramétricos	323
4.2.6.1.4 Análisis de correlación	324
4.2.6.2 Análisis Confirmatorio	326
4.2.6.2.1 Modelo Logit sector hostelería	326
4.2.7 COMPARACIÓN DEL MODELO GLOBAL CON LOS MODELOS SECTORIALES EN M.I	329
4.2.8 MODELO LOGIT GLOBAL CON VARIABLES CUALITATIVAS SECTORIALES	331
4.2.9 ANÁLISIS DE ROBUSTEZ	336
4.3 RESULTADOS CON INFORMACIÓN DE DOS AÑOS ANTES DE LA QUIEBRA (M.II)	340
4.3.1 RESULTADOS MODELO GLOBAL	340
4.3.1.1 Análisis Exploratorio	340
4.3.1.1.1 Análisis Descriptivo	340

	Pág.
4.3.1.1.2 Análisis de Normalidad	341
4.3.1.1.3 Tests no paramétricos	342
4.3.1.1.4 Análisis de correlación	343
4.3.1.2 Análisis Confirmatorio	344
4.3.1.2.1 Modelo Logit Global sin variables cualitativas	345
4.3.2 RESULTADOS MODELO AGRICULTURA	347
4.3.2.1 Análisis Exploratorio	347
4.3.2.1.1 Análisis Descriptivo	347
4.3.2.1.2 Análisis de Normalidad	347
4.3.2.1.3 Tests no paramétricos	348
4.3.2.1.4 Análisis de correlación	349
4.3.2.2 Análisis Confirmatorio	350
4.3.2.2.1 Modelo Logit sector agricultura	352
4.3.3 RESULTADOS MODELO INDUSTRIA	352
4.3.3.1 Análisis Exploratorio	353
4.3.3.1.1 Análisis Descriptivo	353
4.3.3.1.2 Análisis de Normalidad	354
4.3.3.1.3 Tests no paramétricos	355
4.3.3.1.4 Análisis de correlación	355
4.3.3.2 Análisis Confirmatorio	356
4.3.3.2.1 Modelo Logit sector industria	359
4.3.4 RESULTADOS MODELO CONSTRUCCIÓN	361

	Pág.
4.3.4.1 Análisis Exploratorio	361
4.3.4.1.1 Análisis Descriptivo	361
4.3.4.1.2 Análisis de Normalidad	362
4.3.4.1.3 Tests no paramétricos	363
4.3.4.1.4 Análisis de correlación	364
4.3.4.2 Análisis Confirmatorio	366
4.3.4.2.1 Modelo Logit sector construcción	366
4.3.5 RESULTADOS MODELO COMERCIO Y SERVICIOS	368
4.3.5.1 Análisis Exploratorio	368
4.3.5.1.1 Análisis Descriptivo	368
4.3.5.1.2 Análisis de Normalidad	369
4.3.5.1.3 Tests no paramétricos	370
4.3.5.1.4 Análisis de correlación	371
4.3.5.2 Análisis Confirmatorio	373
4.3.5.2.1 Modelo Logit sector comercio y servicios	373
4.3.6 RESULTADOS MODELO HOSTELERÍA	375
4.3.6.1 Análisis Exploratorio	375
4.3.6.1.1 Análisis Descriptivo	375
4.3.6.1.2 Análisis de Normalidad	376
4.3.6.1.3 Tests no paramétricos	377
4.3.6.1.4 Análisis de correlación	378
4.3.6.2 Análisis Confirmatorio	380

	Pág.
4.3.6.2.1 Modelo Logit sector hostelería	380
4.3.7 COMPARACIÓN DEL MODELO GLOBAL CON LOS MODELOS SECTORIALES EN M.II	381
4.3.8 MODELO LOGIT GLOBAL CON VARIABLES CUALITATIVAS SECTORIALES	384
4.3.9 ANÁLISIS DE ROBUSTEZ	389
4.4 ANÁLISIS COMPARATIVO DE RESULTADOS	392
BIBLIOGRAFÍA CAPÍTULO 4	397
DISCUSIÓN DE RESULTADOS	401
CONCLUSIONES	417
BIBLIOGRAFÍA GENERAL	425
ANEXOS	465
ANEXO 1: INDICE DE TABLAS	467
ANEXO 2: INDICE DE GRÁFICOS	471
ANEXO 3: HISTOGRAMAS	473
ANEXO 4: GLOSARIO DE SIGLAS	521

INTRODUCCIÓN

El análisis de la quiebra empresarial ha sido objeto de especial atención en la investigación financiera durante las últimas cinco décadas, siendo muy numerosos y variados los trabajos orientados a determinar los factores que provocan el fallo empresarial. Esta ingente tarea de investigación ha generado una amplia variedad de modelos, soportados, a su vez, en muy diversas metodologías (Dimitras et al., 1996; Bellovary et al., 2007; Sun et al. 2014).

Uno de los caminos inicialmente tomados por la literatura fue el desarrollo de modelos que no habían sido diseñados utilizando una muestra específica compuesta por empresas de un solo sector y que, por tanto, bien podían considerarse como modelos descentrados. Así, con carácter general, los modelos que tradicionalmente se vinieron desarrollando en un principio para la predicción de la insolvencia eran modelos contruidos con muestras formadas por empresas medianas y grandes, pertenecientes al sector industrial y comercial en un sentido amplio.

Podría decirse que el desarrollo más amplio que tuvo la literatura en la construcción de modelos descentrados se debió a la mera necesidad. No es de extrañar que hasta que no se desarrollaron bases de datos más amplias y completas no se llevaran a cabo modelos de predicción de insolvencia centrados en sectores de actividad específicos.

Un análisis de la literatura sobre predicción de insolvencia permite constatar la existencia de un patrón definido por lo que se refiere a la construcción de modelos descentrados frente a modelos centrados, siendo los primeros mucho más numerosos en número que los segundos. En cambio, no es posible identificar ninguna tendencia, ni a favor ni en contra, en el uso de un tipo de modelo u otro, como señala Bellovary et al. (2007), siendo imposible inducir una conclusión definitiva sobre la superioridad de un tipo de modelo frente a otro.

Diversos autores ya afirmaron que las particularidades sectoriales añadían mejoras significativas en la explicación ofrecida por los modelos, pero lo cierto es que esta cuestión no ha sido abordada de manera directa con un diseño *ad-hoc* para tal objetivo.

Una posible explicación, a esta ausencia de un patrón definido en el uso de un tipo u otro de modelos, podría basarse en la utilización de muestras heterogéneas de empresas pertenecientes a diferentes sectores o de muestras centradas en empresas que operaban en el sector industrial. En este sentido, es innegable, como afirman Platt y Platt (1991), que los sectores económicos muestran diferencias significativas con respecto a los factores de producción, ciclos de vida de sus productos, estructura competitiva y modos de distribución, lo que da lugar a diferencias importantes en diversas medidas de salud financiera. Esta idea ya fue intuita por Lincoln (1984), Mensah (1984) y McDonald y Morris (1984),

quienes analizaron la influencia de los efectos sectoriales en los ratios financieros y su aplicación en la predicción del fracaso empresarial.

Sin embargo, es posible que la ausencia de una conclusión práctica sobre la superioridad de un modelo centrado sobre un modelo descentrado o viceversa sea debida a que no se hayan podido comparar, de forma homogénea, un tipo de modelo y otro debido a la disparidad de metodologías, enfoques, bases de datos disponibles, periodos temporales y países, entre otras cuestiones.

Por lo tanto, la existencia de este hueco en la literatura, que no permite dilucidar la superioridad de los modelos descentrados frente a los modelos centrados, es el que nos ha motivado a elaborar esta tesis doctoral con el objetivo de dilucidar la posible superioridad de unos modelos frente a otros a la hora de predecir la insolvencia empresarial.

Para resolver la cuestión de investigación planteada se ha diseñado el presente trabajo conforme a la estructura siguiente. En el capítulo 1 se lleva a cabo un repaso sistemático de la literatura previa sobre predicción de insolvencia, a partir de la cual se puede constatar el desarrollo de las aportaciones realizadas a lo largo de las pasadas décadas. De este primer análisis emerge una imagen con unos perfiles muy definidos, por lo que se refiere a los diferentes caminos seguidos por la

investigación en predicción de insolvencia, con una consideración especial a los métodos utilizados y su perfeccionamiento, a las variables empleadas más relevantes y a la evolución de los resultados obtenidos por los diferentes modelos.

En segundo lugar, y atendiendo al enfoque de estudio adoptado en esta tesis, la literatura se ha analizado diferenciando si se trataba de modelos de predicción de insolvencia descentrados, es decir, contruidos sobre la base de muestras heterogéneas en cuanto a la pertenencia sectorial de las empresas consideradas, o bien si se trataban de modelos centrados por cuanto su elaboración se apoyó en el empleo de empresas pertenecientes a sectores o subsectores de actividad específicos.

El capítulo 1 concluye ofreciendo una taxonomía de los estudios sobre predicción de insolvencia, en la que se ponen de manifiesto las diferentes tipologías de estudios junto con una exposición de los principales argumentos que ponen de relieve el “gap” existente en la literatura acerca de la superioridad de los modelos centrados frente a los modelos descentrados, y que se materializan en las tres hipótesis a contrastar en la presente tesis.

El capítulo 2 desarrolla los métodos principales de clasificación binaria de naturaleza estadística. En especial, se presentan los modelos de regresión logística, que es la metodología empleada en esta tesis. Asimismo, en este

capítulo se dedica una especial atención al desarrollo de los principales criterios de selección de modelos de carácter informacional.

En el capítulo 3 se exponen tanto el proceso de obtención y el tratamiento de las muestras de trabajo como las variables utilizadas y los criterios tenidos en cuenta para la selección de las mismas. Con respecto a las variables empleadas, se señalan las variables que han sido utilizadas con mayor frecuencia en modelos descentrados. Igualmente, se especifican los criterios de elección de las variables que, en los modelos centrados, han resultado ser variables diferenciales y específicas en la explicación de la insolvencia en los cinco sectores económicos considerados en este estudio: agricultura, industria, construcción, comercio y servicios, y hostelería.

El capítulo 4 se dedica al desarrollo de los resultados de la investigación empírica. En el mismo se obtienen los modelos de predicción de la insolvencia, tanto globales como sectoriales, con información correspondiente a 1 y 2 años antes de que ésta se produzca, utilizando para ello las muestras obtenidas específicamente para este trabajo. Asimismo, se lleva a cabo un análisis de robustez de los resultados obtenidos, mediante una validación cruzada de los modelos y a través de la aplicación de los criterios de selección de modelos de Akaike, Schwartz y Hannan-Quinn. Por último, se ofrece un análisis comparativo de los resultados obtenidos.

El trabajo concluye con una discusión sobre los resultados alcanzados, las principales conclusiones obtenidas del mismo y la bibliografía consultada.

PARTE I:

INVESTIGACIÓN DOCUMENTAL **Y METODOLÓGICA**

CAPÍTULO 1:

LA INVESTIGACIÓN PREVIA EN PREDICCIÓN DE INSOLVENCIA

1.1 INTRODUCCIÓN

En este capítulo se realiza un análisis de la literatura previa que focaliza en dos aspectos relacionados con la presente investigación.

Así, en primer lugar, se lleva a cabo un repaso sistemático de la literatura previa sobre predicción de insolvencia, a partir de la cual se puede observar el desarrollo de las aportaciones realizadas a lo largo de las últimas décadas. De este primer análisis se derivan conclusiones relevantes respecto al desarrollo de la literatura, con una consideración especial a los métodos utilizados y su perfeccionamiento, las variables empleadas más relevantes y la evolución de los resultados obtenidos por los diferentes modelos.

En segundo lugar, y atendiendo al enfoque de estudio adoptado en esta tesis, la literatura se ha analizado discriminando si se trataba de modelos de predicción de insolvencia descentrados, es decir, contruidos sobre la base de muestras heterogéneas en cuanto a la pertenencia sectorial de las empresas consideradas, o bien si se trataban de modelos contrados por cuanto su elaboración se sustentó en el empleo de empresas pertenecientes a sectores o subsectores de actividad específicos.

El análisis combinado de ambos enfoques nos ha permitido extraer unas conclusiones específicas en cuanto al patrón seguido por la literatura en su desarrollo, distinguiendo hacia donde se han canalizado los esfuerzos hasta la fecha y hacia donde, en nuestra modesta opinión, deberían canalizarse en adelante.

En resumen, este análisis de la investigación previa es el que nos ha permitido establecer el marco de referencia en el que se inserta la presente investigación y la formulación de las hipótesis que corresponden a la misma.

1.2 LA PREDICCIÓN DE INSOLVENCIA EN LAS EMPRESAS ECONÓMICAS

El análisis de la quiebra empresarial ha sido objeto de especial atención en la investigación financiera durante las últimas cinco décadas. Se han realizado numerosos trabajos de investigación orientados a determinar los factores que provocan la insolvencia empresarial, con una especial incidencia en cómo predecirla antes de que ésta suceda. Esta ingente tarea de investigación ha generado una amplia variedad de modelos, soportados a su vez, en muy diversas metodologías (Dimitras et al., 1996; Bellovary et al., 2007; Sun et al., 2014).

Del universo de empresas que han sido objeto de estudio, se pueden establecer dos categorías principales: las empresas económicas y las empresas financieras. Ambos tipos de empresas presentan notables diferencias en su casuística y en sus características financieras, por lo que han tenido un tratamiento muy diferenciado.

Por lo que se refiere a las empresas económicas, los autores pioneros de los estudios empíricos sobre predicción de insolvencia fueron Beaver (1966) y Altman (1968), quienes aplicaron el análisis discriminante, univariante, en el caso de Beaver y, multidiscriminante, en el caso de Altman.

Los ratios financieros utilizados por Altman en su análisis multidiscriminante fueron aplicados en numerosos estudios de predicción de insolvencia posteriores, no sólo incluyéndolos parcial o totalmente dentro del conjunto de variables, sino incluso empleando dichos ratios en metodologías estadísticas más complejas.

A partir de estos estudios iniciales, la principal cuestión en la literatura, en lo referente a la predicción de insolvencia, no sólo fue determinar qué ratios o variables incluir en los modelos, sino valorar que método estadístico era el más eficaz para predecir la insolvencia. En este sentido, las investigaciones se apoyaron en los avances producidos en la estadística, lo que hizo posible formular modelos con un mayor poder de predicción y una mayor robustez en sus resultados.

Esta es la auténtica razón por la cual la literatura sobre predicción de insolvencia, ante la ausencia de una teoría global explicativa del fenómeno del fracaso empresarial, se ha desarrollado con fuerza de la mano del progreso y la aplicación de diferentes metodologías, tanto de carácter estadístico como de naturaleza computacional. En consecuencia, el desarrollo de la literatura sobre predicción de insolvencia debe de ser contemplada en paralelo con la evolución seguida por las diferentes metodologías empleadas.

Así, un repaso de los métodos empleados a lo largo de estas décadas de investigación nos permite entender la evolución seguida en la literatura². Atendiendo a este criterio,

² Para un estudio de la utilidad de los métodos de evaluación del riesgo de insolvencia en las pymes véase Lamothe, Gómez y Partal (2010).

los primeros trabajos se desarrollaron en torno a los llamados clasificadores individuales puros. Dentro de éstos se encuentran los clasificadores estadísticos que incluyen el análisis individual de variables, el análisis multidiscriminante (MDA) y los modelos de regresión logística (Logit), que están basados en la teoría estadística y son, por regla general, simples y fáciles de usar. El análisis individual de variables fue el primer método usado en la predicción de insolvencia. Ya se mencionó anteriormente a Beaver (1966), que propuso dos métodos de análisis individual de variables, llamados análisis de perfil y modelo discriminante univariable. Mediante el análisis de perfil para cinco años antes de la quiebra, descubrió que los resultados de ratios financieros en dos grupos de empresas eran significativamente diferentes, haciéndose más patente esa diferencia a medida que se acercaba el momento de la quiebra.

De la misma forma, y como se apuntó anteriormente, Altman (1968) fue el primero en usar el MDA para predicción de insolvencia, técnica que pertenece a la categoría del análisis discriminante multivariante. Construyó el famoso modelo "Z-score", que es una función discriminante lineal multivariable con cinco ratios financieros, demostrando que su capacidad predictiva en el año inmediatamente anterior a la quiebra era significativamente mejor que el modelo discriminante de sólo una variable. El MDA pertenece a los análisis de regresión lineal, los cuales asumen que la variable independiente debería presentar una distribución normal multivariable y una igualdad en su matriz de covarianzas.

El modelo Logit de probabilidad lineal usa las funciones logísticas para transformar la variable dependiente de probabilidad de insolvencia en una totalmente continua que

luego es adecuada para un análisis de regresión lineal. Fue Ohlson (1980) quien aplicó originalmente el modelo Logit para describir la relación entre la situación de insolvencia y los ratios financieros, sugiriendo que dicho modelo era comparativamente mejor que el MDA para la predicción de insolvencia. Utilizando esta misma metodología, Zavgren (1985) usó el análisis de factores para seleccionar las variables independientes del modelo Logit. Posteriormente, Tseng y Hu (2010) propusieron un intervalo cuadrático del modelo Logit mediante el uso de un enfoque de programación cuadrática para hacer frente a las variables de respuesta binaria en la regresión logística. El resultado del nuevo modelo Logit mejoró el rendimiento discriminante y aportaba una mayor información a los investigadores.

Más recientemente, Serrano-Cinca y Gutiérrez-Nieto (2013) aplicaron un análisis discriminante de mínimos cuadrados parciales (PLS-DA) para predecir la crisis bancaria americana del periodo 2008-2011, y encontraron que los resultados eran similares a los obtenidos mediante análisis discriminante multivariante (MDA) y las Máquinas de Soporte Vectorial (SVM). El método PLS-DA tenía la ventaja de que no le afectaban los problemas de multicolinealidad.

Dentro de los denominados clasificadores de inteligencia artificial, se incluyen las Redes Neuronales (NN), los Algoritmos de Evolución (EA), el Rough Set (RS), el Razonamiento Basado en Casos (CBR) y las Máquinas de Soporte Vectorial (SVM). Todos ellos fueron también aplicados en la predicción de insolvencia y venían avalados por unos fructíferos resultados en las investigaciones sobre inteligencia artificial y computacional.

Aportaron la ventaja de que no estaban sujetos a los estrictos supuestos requeridos por los métodos estadísticos.

Es a principios de los años 90 cuando las NN comienzan a ser introducidas en las investigaciones de predicción de insolvencia de la mano de Tam (1991) y Tam y Kiang (1992). Un modelo de NN consiste en un grupo interconectado de neuronas artificiales que procesan la información usando un enfoque conexionista para el cálculo. La red neuronal probabilística fue también usada ampliamente en la predicción de insolvencia y empleada en las teorías bayesianas de toma de decisiones para la estimación de la densidad probabilística en espacios de datos (Wu et al., 2008; Yang et al., 1999). Callejón et al. (2013) aplicaron la variante de NN conocida como Perceptrón Multicapa con el fin de optimizar los resultados de los modelos NN que se estaban aplicando en las investigaciones empíricas.

Las ventajas de las NN sobre los métodos estadísticos tradicionales se atribuyen frecuentemente a su alta capacidad de aplicación práctica basada en una estructura de red. Además, las relaciones estadísticas entre variables no necesitan ser tenidas en cuenta en el proceso de construcción del modelo NN. En cambio, comparado con los métodos estadísticos, se necesita una muestra de datos de entrenamiento superior para generar un modelo de red relativamente estable, sin olvidar que demasiado entrenamiento suele conducir a un sobreajuste, disminuyendo la estabilidad de predicción de la muestra apartada para el testeo.

Otro de los métodos de inteligencia artificial relativamente nuevo, que se basa en el principio de minimización del riesgo

estructural en lugar de principio de minimización del riesgo empírico, son las Máquinas de Soporte Vectorial (SVM). Esta técnica es un potente y prometedor clasificador de datos y una herramienta de estimación funcional (Wang et al., 2005) con la que no es fácil caer en el sobreajuste, incluso en muestras relativamente pequeñas. Shin et al. (2005) y Min y Lee (2005) emplearon ambas las SVM para predecir la insolvencia en empresas surcoreanas y llegaron a la conclusión de que este método supera en rendimiento al MDA, Logit y NN. Hui y Sun (2006) y Ding et al. (2008) adoptaron un modelo de SVM para hacer un estudio empírico sobre predicción de insolvencia para una serie de compañías chinas y el resultado arrojó unas conclusiones similares.

Sin embargo, Bose y Pal (2006) compararon el rendimiento de las SVM con el de las NN y MDA, concluyendo que las SVM eran peores. Otros estudios usaron las SVM junto con otras técnicas como las híbridas o las conjuntas, en vez de hacer un uso independiente de las SVM.

Junto a los anteriores, también destaca, como clasificador de inteligencia artificial, los Algoritmos de Evolución (EA). Los EA son algoritmos genéricos basados en la población meta heurística que usan mecanismos inspirados en la evolución biológica, como la reproducción, mutación, recombinación y selección. Varetto (1998) aplicó los Algoritmos Genéticos (GA) para extraer funciones lineales sin restricciones estadísticas y sus correspondientes reglas discriminantes. Sin embargo, el rendimiento de su modelo de predicción de insolvencia no fue tan bueno como el MDA. Shin y Lee (2002) adoptaron los GA para buscar el punto de inflexión de los ratios financieros, con

objeto de extraer las reglas discriminantes cuantitativas en la predicción de insolvencia.

Además, Rafiei et al. (2011) también encontraron que los GA tienen una menor precisión que las NN. Etemadi et al. (2009) estudiaron la aplicación de la Programación Genética (GP) en la predicción de insolvencia para sociedades de Irán, y el test de McNemar demostró que supera al MDA. Por lo tanto, los métodos basados en EA pueden generar reglas más fácilmente interpretables que los modelos de NN o SVM, pero su rendimiento se encuentra más limitado y serían más adecuados para ser combinados con otros algoritmos de clasificación.

El Razonamiento Basado en Casos (CBR) usa el algoritmo de K-vecinos más cercanos, el cual determina o predice los casos objetivos a partir de unos casos similares. La ventaja evidente del CBR es que es fácil de comprender y la precisión de su predicción también es relativamente alta. Comparando CBR, NN y MDA, Jo y Han (1996, 1997) encontraron que no había una diferencia real entre MDA y CBR, y que el CBR funcionaba bien cuando no existían datos suficientes.

Por su parte, Park y Han (2002) aplicaron la ponderación del algoritmo de K-vecinos más cercanos a la predicción de insolvencia y mejoraron significativamente la precisión de la predicción integrando indicadores tanto cuantitativos como cualitativos. Además de lo anterior, Sun y Han (2006) propusieron una ponderación similar basada en la distancia de correlación gris, la cual funciona en el corto plazo de la predicción de insolvencia con uno o dos años de antelación. Por último, y más recientemente, Borrajo et al. (2011) construyeron un sistema multi-agente para los procesos de

control de negocios y predicción de insolvencia, incorporando un sistema CBR al agente del núcleo.

Dimitras et al. (1999) y McKee (2000), respectivamente, aplicaron la teoría del Conjunto Aproximado o Rough Set (RS) para establecer un modelo de predicción de insolvencia en Grecia y Estados Unidos. Este método tenía muchas ventajas, como por ejemplo, reglas de decisión fácilmente comprensibles en lenguaje natural, apoyo a las reglas de decisión, combinación de variables cualitativas y cuantitativas sin necesidad de evaluación estadística de la probabilidad, etc. Sin embargo, diferentes muestras pueden generar diferentes grupos de reglas de decisión. Por lo tanto, el RS tendría las deficiencias de una estructura cambiante y una pobre universalidad.

Friedman et al. (1985) y McKee y Greenstein (2000) adoptaron un método de partición recursiva que generó árboles de decisión (DT) para predicción de la insolvencia. Sun y Li (2008) propusieron un método de minería de datos basado en la inducción de atributos, ganancia de información y DT, y lo aplicaron a una serie de compañías chinas. Gepp et al. (2010) y Li et al. (2010) también demostraron que los DT son mejores que el MDA. Chen (2011) comparó empíricamente los DT con el modelo Logit para empresas de Taiwán y encontró que los DT tienen una mayor precisión que el Logit en el corto plazo (menos de un año), mientras que el rendimiento del Logit es mejor en el largo plazo (alrededor de un año y medio).

Además de los métodos principales anteriormente expuestos, hay algunos otros, por ejemplo, Sarkar y Sriram (2001), Gesel et al. (2006), Sun y Shenoy (2007) y Wu (2011)

aplicaron clasificadores Bayesianos para la predicción de insolvencia. Pendharkar (2002), Cielen et al. (2004) y Premachandra et al. (2011) usaron el análisis envolvente de datos como una herramienta también para la predicción de insolvencia empresarial. Kwak et al. (2011) propusieron un método de programación lineal multi-criterio (MCLP) con datos de Corea, e indicaron que dicho modelo funcionaba tan bien como el MDA o el Logit y fue comparable a los DT y SVM. Ryu y Yue (2005) desarrollaron un nuevo modelo de programación matemática llamado Separación Isotónica (IS), el cual superaba a MDA, NN, DT y RS en un experimento de predicción de insolvencia con datos de Norteamérica.

Durante los últimos años, el grueso de la investigación se ha venido enfocando hacia los métodos híbridos para la predicción de insolvencia empresarial basados en dos o tres algoritmos, siendo los métodos más populares la integración de NN, CBR o SVM con otras técnicas.

Back et al. (1996) y Anandarajan et al. (2001) combinaron NN con GA. Ellos usaron GA para seleccionar las variables de entrada para la NN y establecieron un modelo GA-NN para mejorar adicionalmente el rendimiento. Ravisankar et al. (2010) y Ravisankar y Ravi (2010) también construyeron varios modelos híbridos siguiendo este paradigma. Anteriormente, cada híbrido compuesto por dos de las NN multicapa con alimentación hacia adelante, NN probabilísticas, RS y programación genética, mostraron el mejor rendimiento con un nivel de significación del 10%.

En este último, las características fueron en primer lugar seleccionadas con el estadístico-t, estadístico-f o el método

grupales de manejo de datos (GMDH, una NN relativamente poco explorada), y los clasificadores fueron entrenados con GMDH, NN de propagación y un mapa teórico de resonancia adaptativa borrosa. Como resultado, GMDH-GMDH y el estadístico-t-GMDH superaron a los demás.

Yeh et al. (2010) construyeron un modelo híbrido de dos etapas que integra teoría de RS y SVM, llamado RST-SVM. Los atributos redundantes fueron reducidos mediante RS y luego la insolvencia empresarial se predice con SVM, demostrándose que supera al binomio RST-BNP. Lin et al. (2011) entrenaron un modelo de SVM después de la reducción de características de dimensionalidad, a pesar de las características del algoritmo de mapeo isométrico (ISO-MAP), una de las técnicas de reducción de la dimensionalidad más desarrollada. Esto produjo un mejor rendimiento cuando se compara con técnicas híbridas de PCA-SVM.

Ahn y Kim (2009) presentaron un modelo de predicción de insolvencia que combinaba GA y CBR, donde el GA fue usado para optimizar la ponderación de las características y la selección de ejemplos de forma simultánea. Este híbrido obtuvo una precisión de predicción más alta que el CBR convencional, pero necesita mucho más tiempo de modelaje y recursos de equipo.

Li et al. (2010) únicamente usaron el GA para seleccionar las características para el CBR y aplicaron el algoritmo híbrido a la predicción de insolvencia. Cho et al. (2010) usaron los DT para dibujar variables explicativas aplicando la distancia de Mahalanobis, una forma de determinar la similitud entre dos variables aleatorias multidimensionales teniendo en cuenta

también la correlación entre dichas variables, y superó al método CBR usando la distancia Euclídea, Logit y NN.

Es evidente, pues, que tanto los métodos estadísticos como los computacionales han condicionado el desarrollo de la literatura sobre predicción de insolvencia. Sin embargo, existen otros criterios relevantes para comprender la evolución seguida por la investigación como son los referentes al número de ratios y variables empleadas y a la evolución del grado de acierto en los resultados obtenidos.

Si atendemos al número de variables consideradas en los modelos, siguiendo a Bellovary et al. (2007), el número de factores o ratios considerados en la literatura ha oscilado entre 1 y 57. Un número total de 752 variables diferentes fueron utilizados en los estudios sobre predicción de insolvencia, siendo 674 las variables o factores que fueron utilizadas en uno o dos estudios. Las variables más comúnmente utilizadas en múltiples estudios fueron el ratio Beneficio después de Impuestos/Total Activo, que se fue incluido en 54 estudios. El segundo factor más frecuentemente utilizado fue el ratio de Activo corriente/Pasivo corriente, que se utilizó en un total de 51 estudios. Se constata, además, que ha habido una cierta fluctuación en relación con el número de ratios empleados en los estudios sobre predicción de insolvencia durante los últimos 40 años, si bien la media se ha mantenido relativamente constante en el entorno de 8 a 10 variables.

Las conclusiones obtenidas de la literatura previa resaltan la importancia de, al menos, dos factores en la consecución del equilibrio financiero. De una parte, la adecuada correspondencia entre las fuentes de financiación y la naturaleza de las

inversiones realizadas, Altman (1968) y Zhang et al. (1999), y de otra, la capacidad de la empresa para devolver su deuda utilizando los fondos que se generan con la actividad, Bernstein (1984), Baldwin y Glezen (1992), García y Fernández (1992), y Mulford y Comiskey (2005). Estas conclusiones posibilitan clasificar a las compañías en solventes o insolventes, aunque también pueden utilizarse, no sólo dos categorías de empresas, sino distintos procesos de fracaso empresarial, Jones y Hensher (2007) y Arquero et al. (2009).

Las capacidades predictivas de los modelos de predicción de insolvencia han ido cambiando en función del tiempo y del método empleado. Así, la literatura sobre predicción de insolvencia sigue refiriéndose a los errores tipo I (clasificar como insolvente a una empresa solvente, o falso positivo) y tipo II (clasificar como solvente a una empresa insolvente, o falso negativo), si se tienen en cuenta los resultados obtenidos por estos estudios. Generalmente, se ha aceptado que los errores tipo I son más onerosos que los errores tipo II por diversas razones, entre las que se incluyen la pérdida de negocio y clientes, el daño a la reputación de la empresa y los costes potenciales de los procedimientos judiciales.

Según Bellovary et al. (2007), en la década de los sesenta del siglo pasado el rango de precisión de los modelos se situó entre el 79% y el 92%. Durante esta década, los modelos predominantes estaban basados en el Análisis Discriminante Univariante (Beaver, 1966).

En la década de los setenta el rango de acierto en los modelos se amplió hasta situarse entre un 56% y un 100%. En esta década los modelos predominantes estaban basados en

MDA (Meyer y Pifer, 1970; Edmister, 1972; Santomero y Vinso, 1977).

En los ochenta, el rango de acierto en los modelos se amplió aún más, pasando del 20% al 100%. Ya en esta década empezaron a utilizarse otros métodos además del MDA, como las NN o los Algoritmos Recursivos Particionados. En este sentido destacan los trabajos de Marais (1980), Betts y Behoul (1982), El Hennaway y Morris (1983), Izan (1984), Takahashi et al. (1984), Friedman et al. (1985) y Messier y Hansen (1988).

En la década de los noventa, el rango de precisión en la predicción de los modelos se situó entre el 27% y el 100%, siendo los métodos predominantes NN y Sumas Acumulativas. Destacan los trabajos de Guan (1993), Tsukuda y Baba (1994), El-Temtamy (1995), Koundinya y Puri (1992) y Theodossiou (1993).

Además, la mayor parte de los trabajos realizados han estimado modelos de predicción utilizando información correspondiente a un año antes de la quiebra empresarial, y obteniéndose, de media, un nivel de predicción del 81%.

Son pocos los que han realizado predicciones de quiebra en un mayor plazo, destacando entre los que sí lo realizaron Altman (1968), Diamond (1976), Ohlson (1980), Skogsviki (1980), Coats y Fant (1993), Atiya (2001), Wu et al. (2007), y Lin (2009). De los resultados obtenidos se desprende que la capacidad de predicción de los modelos disminuye considerablemente cuando se utiliza información financiera de más de un año previo a la situación de quiebra de las empresas,

con aciertos que oscilan entre el 72,0% (Altman, 1968) y el 95,5% (Ohlson, 1980). Estos porcentajes de acierto disminuyen aún más en el caso de aquellos estudios que han estimado la quiebra empresarial 3 años antes, oscilando los niveles de acierto entre el 48% (Altman 1968) y el 86,2% (Coats y Fant 1993), con la excepción de Wu et al. (2007) que llegaron a alcanzar un 100% en sus modelos de NN y GA tanto para 1 como para 2 y 3 años antes de la entrada en quiebra.

En términos generales, parece que a medida que los métodos evolucionaban eran capaces de predecir con mucha mayor precisión. Sin embargo, si consideramos los intervalos de predicción, se observa que el valor de los rangos inferiores ha ido disminuyendo con el tiempo evidenciándose, de esta forma, una dilatación del rango de predicción.

Esto parece sugerir que los métodos más modernos no siempre garantizan los mejores resultados. De hecho, en numerosos estudios, MDA y NN continúan proporcionando las mayores tasas de acierto junto con los menores rangos de predicción. Y si tuviéramos que destacar una metodología en concreto, NN resultan ser el método que ofrece el rango de resultados más reducido junto con el mayor porcentaje de acierto, oscilando desde un 71% al 100% de acierto.

1.3 MODELOS GLOBALES DE PREDICCIÓN DE INSOLVENCIA

Otro término utilizado en la literatura es el que se refiere a modelos globales como aquellos que consideran a las empresas de todo un país o región.

En este sentido, los trabajos que incorporan esta particularidad o característica son de muy diversa naturaleza. En la mayoría de los casos son estudios que se han centrado en la predicción de la insolvencia para empresas de un solo país. Por ejemplo, Serrano-Cinca (1997) en el caso de empresas españolas, o Charalambous et al. (2000) con empresas estadounidenses.

Una segunda tipología de trabajos con particularidad global son los elaborados teniendo como referencia a empresas de distintos países. Así, Tsai (2009) tuvo en cuenta las empresas de Australia, Alemania y Japón, o Chen et al. (2011) las empresas polacas y australianas. Zhou (2013) elaboró un modelo con empresas estadounidenses y japonesas, empleando una batería de métodos como el Análisis Discriminante Lineal, Logit, DA, NN y SVM.

Pero, además de incorporar esta característica global, los enfoques adoptados en este tipo de investigaciones son diversos. Así, Pindado et al. (2008) formulan un modelo de probabilidad de impago tomando como muestra de referencia empresas de Estados Unidos y del resto de países del G-7. Chauhan et al. (2009) aplicaron una NN Ondícula con una muestra de bancos estadounidenses, turcos y españoles. Ravisankar y Ravi (2010) utilizaron una muestra muy diversa, con bancos españoles, turcos, estadounidenses y británicos, obteniendo una alta precisión (por encima del 90%) y empleando el método Grupo de Gestión de Datos (GMDH). Evans y Borders (2014) propusieron una comparación entre los códigos (legislativos) de quiebra de los diferentes países.

Por su parte, Aggarwal y Goodell (2014) prestaron especial atención a las diferencias entre países en lo referente al acceso a la financiación y a su influencia en las empresas con dificultades financieras. Maghyreh y Awartani (2014) aplicaron un modelo logit multi-periodo para bancos de países del Consejo de Cooperación del Golfo durante el periodo 2000-2009. Tsai et al. (2014) aplicaron un Perceptrón Multicapas comparándolo con las SVM y DT para una muestra formada por un conjunto de empresas alemanas, japonesas y australianas. Betz et al. (2014) estudiaron las dificultades financieras en bancos europeos en los 3 años previos a la quiebra con un modelo Logit.

Por otro lado, el artículo de Korol (2013) incorpora el enfoque de la quiebra y hace una comparación entre dos regiones. Realiza un modelo de riesgo de quiebra basado en la idea de sistema de "Advertencia Temprana" usando la estructura del modelo de Altman con países de América Latina (México, Argentina, Brasil, Chile y Perú) y de Europa Central (Polonia). Utiliza MDA, DT y NN como métodos, y 14 ratios financieros como posibles variables. Llegó a la conclusión de que los errores de tipo I eran más altos en las empresas latinoamericanas que en las empresas europeas. Los resultados mostraron que el modelo DT tiene una mayor eficacia que el modelo NN en ambas muestras. Pero hay que advertir que este trabajo se centró en hacer una comparación entre los modelos de América Latina y Europa Central y no se ocupó de cerrar la principal brecha de la investigación que es la construcción de modelos específicos para cada región de importancia en el mundo. Tan sólo alcanza a comparar métodos y resultados entre dos regiones. Además, la muestra de Europa Central que emplea no es del todo suficiente, al incluir únicamente

empresas polacas. Tampoco hace referencia a la composición industrial de sus muestras ni a la justificación de los ratios financieros elegidos.

Finalmente, el trabajo de Platt y Platt (2008) se centró en estudiar las dificultades financieras de empresas del sector industrial en tres regiones del mundo (Estados Unidos, Europa y Asia), adoptando para ello un concepto amplio del concepto de dificultad financiera. Las hipótesis formuladas en el trabajo pueden considerarse como muy poco restrictivas, centrándose en la elaboración de modelos Logit y obviando realizar alguna prueba con cualquier modelo NN. Igualmente, el tratamiento de los datos sólo tuvo en consideración muestras de entrenamiento, sin chequear el modelo en muestras de testeo. Y tampoco se situaron algunos años previos a la situación de dificultad financiera, para intentar anticiparse al problema.

1.4 MODELOS DESCENTRADOS Y MODELOS CENTRADOS

Como ya apuntamos anteriormente, los estudios sobre predicción de insolvencia se han aplicado tanto en el terreno de la empresa financiera como en el de la económica. Dentro del ámbito de éstas, a partir de los primeros estudios de Beaver y Altman, la literatura se fue desarrollando en múltiples direcciones.

Uno de los caminos tomados por la literatura fue el desarrollo, por una parte, de modelos que no habían sido diseñados utilizando una muestra de un solo sector y que, por tanto, bien pudieran denominarse modelos descentrados. Solo, posteriormente, la literatura se fue desarrollando mediante la aparición de modelos centrados en la predicción de la

insolvencia en sectores económicos concretos, en la medida en que habían sido contruidos mediante muestras de sectores económicos específicos.

Así, con carácter general, los modelos que se vinieron desarrollando tradicionalmente para la predicción de la insolvencia eran modelos contruidos con muestras formadas por empresas medianas y grandes, pertenecientes a los sectores industrial y comercial en un sentido amplio. Dentro de estos modelos descentrados podríamos destacar, como más relevantes, los trabajos de Beaver (1966), Altman (1968), Deakin (1972), Wilcox (1973), Libby (1975), Moyer (1977), Ketz (1978), Norton y Smith (1979), Casey (1980), Ohlson (1980), Aharony et al. (1980), Zimmer (1980), Zmijewski (1984), Gentry et al. (1985), Levitan y Knoblett (1985), Casey y Bartczak (1985), Gombola et al. (1987), Karels y Prakash (1987), Hing-Ling Lau (1987), Aziz et al. (1988), Hopwood et al. (1989), Aziz y Lawson (1989), Odom y Sharda (1990), Koh y Killough (1990), Gilbert et al. (1990), Coats y Fant (1992), Fletcher y Goss (1993), Theodossiou (1993), Johnsen y Melicher (1994), Altman (1994), Wilson y Sharda (1994), Ward (1994), Boritz y Kennedy (1995), McKee (1995), Leshno y Spector (1996), Serrano-Cinca (1996), Chang-Lee et al. (1996), Kahya y Theodossiou (1999), Charalambous (2000), Nasir et al. (2000), Shumway (2001), Yi-Chung et al. (2005), Pindado et al. (2008), Tzong-Huei Lin (2009), Muller et al. (2009), Shuk-Wern Ong et al. (2011), Chen et al. (2011), Kwak y Gang Kou (2012), Sangjae y Wu (2013).

En consecuencia, teniendo en cuenta este planteamiento general, el desarrollo de modelos que emplean, como muestra, empresas de sectores o subsectores más específicos se

consideran modelos centrados, en contraposición con los modelos descentrados antes reseñados.

Precisamente, por lo que se refiere a los modelos centrados, los más populares inicialmente en la literatura fueron los que se desarrollaron para predecir la insolvencia en las instituciones de crédito. Entre estos, caben destacar los modelos de Meyer y Pifer (1970), Sinkey (1975), Hanweck (1977), Martin (1977); Santomero y Vinso (1977), Pettway y Sinkey (1980), Rose y Kolari (1985), Lane et al. (1986), Pantalone y Platt (1987a, 1987b), Bell et al. (1990); Espahbodi (1991), Tam (1991), Salchenberger et al. (1992), Tam y Kiang (1992), Martín del Brío y Serrano-Cinca (1995), Henebry (1996), Alam et al. (2000).

Un segundo tipo popular de modelos centrados fueron aquellos que se desarrollaron para la predicción de la insolvencia en empresas industriales. Dentro de esta tipología cabría destacar a Taffler (1974, 1977), Diamond (1976), Tisshaw (1976), Mensah (1983), Appetiti (1984), Tsukuda y Baba (1994), Alici (1996), Sung et al. (1999), Zhang et al. (1999), Goover (2003), Hu y Tseng (2005), Habilollah y Saleh (2009), Xu y Zhang (2009), Jae Kwon Bae (2012), Callejón et al. (2013), Zhang et al. (2013), Grünberg y Lukason (2014) y Bartoloni y Baussola (2014), entre otros.

Más recientemente, se fueron desarrollando modelos centrados en la predicción de la insolvencia en sectores económicos más específicos como el sector de la Agricultura (Susicky, 2011), Construcción (Mínguez Conde, 2006), Hostelería (Gao, 1999), Sector Informático (Shah y Murtaza,

2000), Casinos y Juego (Patterson, 2001) y Empresas de Internet (Wang, 2004), por reseñar algunos de ellos.

1.4.1 MODELOS DESCENTRADOS

Como ya se ha expuesto, los modelos descentrados son aquellos que fueron elaborados mediante muestras de carácter heterogéneo. Es decir, muestras compuestas por empresas que no pertenecían a un solo sector económico.

De alguna forma, el desarrollo más amplio de la literatura en la construcción de modelos descentrados se debió a la mera necesidad. La disponibilidad de fuentes de información y registros válidos que permitieran un tamaño muestral suficiente con el que producir inferencias válidas fue un factor limitador para el desarrollo de modelos más centrados o específicos.

Por ello, no es de extrañar que hasta que no se desarrollaron bases de datos más amplias y completas no se llevaran a cabo modelos de predicción de insolvencia centrados en sectores de actividad específicos.

En este apartado realizaremos un análisis pormenorizado de los modelos de predicción de insolvencia descentrados.

Beaver (1966), en su trabajo pionero, afirma que su objetivo es retar el statu quo existente con respecto al uso de los ratios financieros como predictores de la insolvencia. Para ello, se sirvió de la base de datos de Moody's Industrial Manual, que contenía la información financiera publicada de las empresas cotizadas en Estados Unidos, y seleccionó a 79 empresas industriales (una vez excluidas las empresas de

servicios públicos, transportes y de servicios financieros) que hubieran quebrado durante el periodo 1954 a 1964, agrupándolas de acuerdo al sector y al tamaño de sus activos, y emparejándolas con otras 79 empresas sanas. Los datos financieros de las empresas quebradas se obtuvieron para los 5 años previos a la quiebra, empleando 29 variables en total.

Mediante una comparación de los valores medios de las variables detectó que la diferencia entre las medias de los ratios se acrecentaba conforme las empresas quebradas se acercaban al momento de la quiebra. Mediante un test de clasificación dicotómico comprobó que el ratio Cash Flow/Total Deuda era el ratio que mayor capacidad predictiva presentaba, mostrando en el año previo a la quiebra un error del 13%, y del 22% en el segundo año previo. A continuación le seguía como mejor predictor el ratio Beneficio Neto/Total Activo, siendo el tercer mejor predictor el ratio Total Deuda/Total Activo. En cambio, Beaver señaló que los ratios de liquidez eran los que peor se comportaban a la hora de predecir la insolvencia, exhibiendo un error global del 13% en el año previo a la quiebra y del 22% en el quinto año previo al fallo empresarial.

Beaver aconseja precaución en el uso de los ratios ya que, en su opinión, no todos tienen la misma capacidad predictiva ni son capaces de discriminar de igual forma las empresas sanas de las quebradas, teniendo en cuenta, además, que siempre es más fácil clasificar las empresas sanas que las quebradas.

Altman (1968), considerado el trabajo clásico por excelencia, se plantea valorar la capacidad del análisis de los ratios como herramienta analítica, usando la predicción de la quiebra empresarial como un caso ilustrativo. En concreto,

emplea el Análisis Discriminante como metodología y lo aplica a un conjunto de empresas industriales en sentido amplio. Los ratios empleados fueron Fondo de Maniobra/Total Activo, Beneficios Retenidos/Total Activo, EBIT/Total Activo, Valor de mercado de los Fondos Propios/Valor de mercado de la Deuda Total, y Total Ventas/Total Activo.

La muestra inicial estaba compuesta por 66 empresas (33 empresas sanas y 33 empresas quebradas) durante el periodo 1946-1965. La clasificación de este trabajo como un modelo descentrado, a pesar de haberse considerado tradicionalmente un modelo industrial, se debe a que la muestra empleada no es un grupo totalmente homogéneo debido a diferencias de tamaño y sector. No se tuvo en cuenta ni las empresas pequeñas ni las muy grandes.

El modelo desarrollado por Altman se mostró sumamente eficaz para predecir la insolvencia un año antes de la quiebra, arrojando un 95% de acierto, un error tipo I del 6% y un error tipo II del 3%. Sin embargo, la capacidad del modelo disminuye en la predicción a dos años, donde acierta en el 83% de los casos, disparándose el error tipo I al 28%. En la muestra de testeo, el porcentaje sigue siendo muy bueno un año antes de la quiebra con un 96% de acierto global. No obstante, su eficacia vuelve a decaer rápidamente a partir del segundo, tercero, cuarto y quinto año (72%, 48%, 29% y 36%, respectivamente).

Altman, con su trabajo, trasciende al realizado por Beaver anteriormente, al presentar un marco multivariante de análisis de ratios con resultados muy positivos.

Posteriormente, Deakin (1972) propuso un modelo alternativo a los desarrollados por Beaver (1966) y Altman (1968) para predecir el fracaso empresarial. Para ello, replicó el estudio de Beaver usando los mismos ratios que él usó para, a continuación, buscar una combinación lineal de estos 14 ratios que mejor predijeran el fracaso empresarial en cada uno de los 5 años previos a la quiebra.

Partiendo de la muestra utilizada por Beaver de 79 empresas quebradas en el periodo 1954-1964, junto con otras 79 empresas sanas de aproximadamente el mismo tamaño e industria, Deakin introduce una diferencia importante: utiliza una muestra de 32 empresas quebradas durante el periodo 1964-1970. Consideró quebradas aquellas empresas que experimentaron el impago, la insolvencia y/o fueron liquidadas para el beneficio de sus acreedores. Cada una de estas empresas quebradas se emparejó con una empresa sana en función de la industria de pertenencia, el año del cual se obtuvo la información y el tamaño de acuerdo con el volumen de activo.

Los resultados obtenidos por Deakin confirman los anteriormente obtenidos por Beaver. Sin embargo, observó que, a partir del tercer año previo a la quiebra, decae la capacidad predictiva del ratio Tesorería/Ventas, a diferencia de lo que se observa con los demás, en los que existe una elevada concordancia con el trabajo de Beaver. Prácticamente todas las variables contribuyeron significativamente a la capacidad discriminatoria de la función lineal. Por su parte, los errores de clasificación promediaron un 3% (1º año), 4,5% (2º y 3º año) y 21% (4º año) y 17% (5º año). A su vez, el modelo se validó en una muestra independiente de 11 empresas quebradas y 23

empresas sanas seleccionadas al azar entre 1963 y 1964, arrojando unas tasas de error del 22%, 6%, 12%, 23% y 15%, respectivamente.

Wilcox (1973) también se propuso revisar el trabajo original de Beaver. Para ello validó la utilidad predictiva de un estadístico derivado de un proceso binomial, mostrando que se produce una mejora sustancial de los resultados sobre los ratios de Beaver para, al menos, cuatro años antes de la quiebra.

La muestra utilizada se conformó con 52 empresas quebradas en el periodo 1949-1971 y se seleccionaron tanto empresas sanas como quebradas que tuvieran entre 5 y 9 años de información financiera.

Los resultados mostraron una importante tasa de aciertos, con un 94% el año previo a la quiebra hasta un 76% el quinto año antes del fallo empresarial.

Libby (1975) presentó un enfoque novedoso con respecto a los anteriores trabajos al afirmar que la utilidad de la información contable es una función de la capacidad predictiva de la información y de la capacidad de los usuarios de la misma para interpretar esos datos.

Presentó los resultados de un estudio de campo, aplicado a la actividad de las entidades financieras, diseñado para evaluar el poder predictivo de la información proporcionada por los ratios y la capacidad de los directivos de préstamos y riesgos para evaluar dicha información en un contexto de fallo empresarial. Asimismo, el análisis también evaluó la consistencia en la interpretación de los ratios de, por una parte,

los ejecutivos de riesgos y préstamos y, de otra, entre los propios ejecutivos.

Así, el trabajo evidenció cómo una pequeña selección de ratios permitió a los responsables de préstamos, con una amplia variedad en sus procedencias, hacer unas predicciones muy ajustadas y fiables sobre el fallo empresarial.

Moyer (1977) trató de solventar algunas de las críticas recibidas por el modelo de Altman. Moyer reexaminó la capacidad predictiva de este modelo sobre la base de una nueva muestra de empresas durante el periodo 1965-1975. Los parámetros del modelo de Altman fueron recalculados a partir de esta nueva muestra, llegando a la conclusión de que la capacidad predictiva del modelo de Altman es muy modesta cuando se aplica a una muestra diferente en un periodo temporal también diferente.

Ketz (1978) parte de la falta de evidencia para saber si el ajuste de los estados financieros, teniendo en cuenta el cambio en el nivel de precios, es ventajoso o no. Por ello, se propuso como objetivo proporcionar evidencia empírica sobre este particular (estados financieros ajustados al nivel de precios frente a estados financieros basados en el coste histórico). Aplicó ambos enfoques para predecir la insolvencia y comprobar cuál de ambas aproximaciones era la mejor.

Comprobó que los ratios calculados a partir de estados financieros que son ajustados para tener en cuenta las variaciones en los precios obtuvieron un mejor resultado, en términos de precisión en la predicción de la insolvencia,

apoyándose en la idea de que los ajustes de los estados financieros a las variaciones en los precios son de utilidad.

Norton y Smith (1979) siguen en la misma línea que Ketz (1978). Tratan de comparar la predicción de la insolvencia basada en ratios calculados a partir de estados financieros que tienen en cuenta las variaciones de los precios (GPL) frente a ratios calculados en estados financieros que usan el criterio del coste histórico (HC). Sin embargo, sus resultados no coinciden con los de Ketz (1978) al mostrarse que, a pesar de las diferencias existentes entre ambos ratios, no queda alterada la capacidad de predicción de la insolvencia por parte de unos y de otros.

El estudio de Libby (1975) afirmaba que los ejecutivos de préstamo de las entidades financieras usando ratios contables eran capaces de predecir la quiebra empresarial con un 74% de precisión. El trabajo de Casey (1980) surge motivado por esta afirmación y por el deseo de replicar de forma más precisa el entorno habitual de toma de decisiones de los ejecutivos responsables de otorgar préstamos. Presentó los resultados de un cuestionario que replicaba el estudio de Libby (1975) bajo algunas condiciones diferentes.

En el estudio de Casey, 46 responsables de préstamos predijeron cuántas, de un total de 30 empresas, irían a la quiebra en el periodo de tres años. Para cada sujeto analizado se computaron tres resultados predictivos (precisión global, precisión para empresas fallidas y precisión para empresas sanas). De la misma forma, se examinó el grado de coincidencia obtenido entre los sujetos. Los resultados mostrados por este estudio y el de Libby (1975) fueron similares por lo que

respecta al grado de acuerdo entre individuos. Sin embargo, encontraron dos ámbitos de discrepancia: (1) Los resultados predictivos de los individuos fueron inferiores debido a los peores resultados obtenidos por las empresas quebradas; (2) Las diferencias individuales, en cuanto al estilo de procesar la información y el nivel de confianza de los individuos, podrían ser determinantes a la hora de explicar la variabilidad de los resultados.

El trabajo de Ohlson (1980) es pionero en presentar una estimación probabilística del fracaso empresarial mediante el uso de un modelo Logit de probabilidad condicional. Ohlson identifica cuatro factores principales, estadísticamente significativos, que afectan a la probabilidad de quiebra en el plazo de un año. Estos factores son: (1) El tamaño de la compañía; (2) Un indicador de su estructura financiera; (3) Un indicador de resultados y (4) Un indicador de liquidez.

La muestra empleada era de los años 1970-1976 y estaba compuesta por 105 empresas quebradas y 2.058 empresas sanas (cotizadas e industriales), obteniéndose tres años de información financiera para las empresas quebradas. Una peculiaridad del estudio de Ohlson es que los datos de las empresas quebradas fueron tomados del estado 10-K, que tenía la ventaja de que permitía conocer el momento de su publicación, con lo que se podía saber si la empresa había entrado en quiebra antes o después de su publicación. Para Ohlson este es un aspecto fundamental ya que se evitaría "sobre-predecir" las empresas quebradas.

Elaboró varios modelos. El primer modelo de Ohlson predecía correctamente la quiebra en el 96,12% de los casos en

el plazo de un año. El segundo modelo que elabora predecía correctamente el 95,55% de los casos en el plazo de dos años, y siempre y cuando la empresa no quebrara al año siguiente. Finalmente, el tercer modelo propuesto por Ohlson predecía acertadamente el 92,84% de los casos en el plazo de tres años.

Los resultados revelaron que hay cuatro factores, estadísticamente significativos, que determinan la probabilidad de quiebra: (1) El tamaño; (2) La estructura financiera expresada a través del ratio Total Pasivos/Total Activos; (3) Una medida de rendimiento económico como es el ratio Beneficio Neto/Total Activos y (4) Una medida de liquidez como es el ratio Fondo de Maniobra/Total Activos.

Hay que referir que, en los tres modelos de Ohlson, el tamaño emergió como un predictor importante.

Según Aharony et al. (1980), a pesar del interés y del esfuerzo dedicado a estudiar la predicción de la insolvencia hasta la fecha, los modelos basados en ratios contables tenían muy poca o ninguna base teórica. Principalmente, lo que se había realizado hasta entonces era combinar diversos tipos de ratios contables de diversa forma y en diversos modelos dando un resultado determinado. Consideran que los datos provenientes del mercado habían recibido muy poca atención hasta ese momento. Los autores creían que los datos de mercado podían proporcionar una base teórica útil a la hora de predecir la insolvencia empresarial en el sentido de que las expectativas de los inversores se verían reflejadas en las medidas de riesgo-rendimiento. Por eso, este trabajo se propuso comparar las características de las empresas sanas y de las empresas quebradas antes de entrar en quiebra

con diversos indicadores de riesgo-rendimiento derivados del CAPM (Capital Asset Pricing Model).

Se observó que ciertas medidas de riesgo basadas en datos de mercado mostraban un comportamiento muy diferenciado entre ambos tipos de empresas. En concreto, la varianza total y la varianza específica de la empresa se comportaban de manera muy diferenciada, según el tipo de empresa, en un periodo de hasta cuatro años previos a la quiebra.

Zimmer (1980) parte del trabajo de Libby (1975) en el que se evaluó la capacidad predictiva de los ejecutivos de préstamos. Consideró la tarea adicional de que los ejecutivos de préstamos evaluaran cada año la probabilidad de insolvencia de las empresas sobre la base de una serie de ratios contables. Asimismo, el trabajo profundizó sobre diversas cuestiones: (a) Se evaluó la capacidad predictiva de los directivos de préstamos, a través del número de predicciones acertadas; (b) Se investigó la relación entre la confianza en la predicción y los resultados obtenidos; (c) Las diferencias obtenidas en las predicciones se analizaron en función de la edad, experiencia en el cargo y cualificación; (d) El grado de consenso también fue examinado; (e) Se comparó el grado de éxito entre las predicciones del consenso y la de los individuos y (f) Los resultados fueron replicados empleando a estudiantes.

El estudio puso en evidencia que los ratios contables convencionales eran útiles para los ejecutivos de préstamos a la hora de predecir la insolvencia. En general, las predicciones de estos individuos fueron muy acertadas y el grado de acuerdo entre ellos fue muy alto, siendo estos resultados consistentes

con los obtenidos por Libby (1975) anteriormente. Sin embargo, los resultados fueron inconsistentes con los resultados obtenidos en su momento por Casey (1980). Hay que precisar que los individuos participantes en este estudio no fueron elegidos al azar ni la información que les fue suministrada era la misma con la que normalmente evaluaban los riesgos financieros.

El trabajo de Zmijewski (1984) analiza dos sesgos que se suelen producir en los estudios de predicción de insolvencia. El primer sesgo, el sesgo de elección de la muestra, se produce cuando primero se observa la variable dependiente y, a continuación, se selecciona una muestra basada en ese conocimiento. Es decir, la probabilidad de pertenecer a la muestra depende directamente de los atributos de la variable dependiente.

El segundo sesgo, el sesgo en la selección de la muestra, se produce cuando sólo se utilizan observaciones con datos completos para estimar el modelo. Ambos sesgos producen unos parámetros asintóticamente sesgados y unas estimaciones sesgadas igualmente.

Con respecto al sesgo de elección de la muestra, lo examinó comprobando las estimaciones probit para un modelo de estrés financiero frente a un modelo probit ajustado, utilizando diversas muestras para así tener una idea de cómo aumentaba el sesgo. Los resultados manifestaron la existencia clara de un sesgo para las muestras cuando el modelo probit no era ajustado. Por otra parte, este sesgo iba disminuyendo a medida que las muestras recogían una proporción de empresas quebradas/sanas más ajustada a la realidad. Sin embargo, este

sesgo no afectaba a los porcentajes de clasificación de las muestras empleadas.

Con respecto al sesgo en la selección de la muestra, comparó las estimaciones de un modelo probit con datos completos frente a un modelo en el que estimó la probabilidad de que las estimaciones incorporasen datos completos. Se comprobó la existencia de un sesgo, si bien tampoco afectó al porcentaje de clasificación final obtenido.

Hasta este momento, los modelos de predicción de insolvencia vinieron empleando distintos tipos de ratios contables o predictores que ayudaran a mejorar los resultados de la clasificación entre empresas sanas y empresas quebradas. Gentry et al. (1985) propusieron que la adición de diversos componentes o ratios que implicaban movimientos de cash flow a los ratios tradicionales, intuyendo que podrían ayudar a mejorar la precisión de las predicciones.

En concreto, afirmaron que dividendos, inversiones y diversos componentes de cuentas de activo corriente podrían aportar una información muy valiosa acerca de la salud financiera de la empresa. Una de las contribuciones de este trabajo radicaba en que los componentes de salida del cash flow eran más importantes que los componentes de entrada a la hora de explicar el fallo empresarial, además de constatarse que la adición de elementos y componentes de cash flow a los ratios tradicionales mejoraba significativamente la calidad de la predicción.

A mediados de los años ochenta surgió el interés por relacionar la predicción de la insolvencia con el trabajo de

auditoría. Así, Levitan y Knoblett (1985) examinaron las variables de los estados financieros que eran más útiles a la hora de predecir la insolvencia, así como las más útiles a la hora de emitir una excepción de "going-concern" en el informe de auditoría.

Construyeron un modelo que usó con éxito ratios financieros para identificar a las empresas que se dirigían a la quiebra. Del estudio emergieron dos implicaciones relevantes: (1) A la hora de emitir excepciones los auditores usan otros criterios que los que se emplearían a la hora de predecir la insolvencia exclusivamente; (2) La emisión de una salvedad implica un asunto grave, en tanto que no emitirla no lleva a ninguna circunstancia definitiva.

Casey y Bartczak (1985) continuaron la línea de trabajos sobre la influencia de las variables de cash flow en la calidad de la predicción de la insolvencia. Trataron de valorar si los datos de cash flow y otros relacionados podrían conducir a unas predicciones más afinadas de quiebra empresarial. Para ello seleccionaron 60 empresas que presentaron declaración de quiebra durante el periodo 1971-1982, extraídas de un conjunto de empresas cotizadas incluidas en Dun & Bradstreet y Wall Street Journal. El criterio para su inclusión fue el de empresas cotizadas que tuvieran al menos 6 años de estados financieros previos a la quiebra. En paralelo, una muestra de 250 empresas sanas se extrajo de Compustat Industrial Tape. Las empresas sanas se eligieron para ser emparejadas con las quebradas de acuerdo con el subsector industrial. El ratio de empresas quebradas por empresas sanas era de 1:4.

Mediante MDA y Logit llegaron a la conclusión de que los datos de cash flow operativo no aportaban un efecto incremental a la capacidad de predicción de los modelos. Estos resultados estaban en desacuerdo con los obtenidos anteriormente por Gentry et al. (1985).

Gombola et al. (1987) continuaron la línea de investigación sobre el papel de las variables de cash flow en la predicción de insolvencia.

Sobre esta cuestión existían dos líneas de investigación principales: a) Por un lado, los que consideraban al cash flow de las operaciones (CFFO) como ingresos más depreciación y trataban de averiguar su impacto en la predicción de insolvencia, y b) Aquellos trabajos que estudiaban la eficacia del CFFO por sí mismo, donde la evidencia obtenida era mixta. Los autores sospechaban que el factor tiempo podía tener un impacto específico a la hora de determinar la influencia de dichas variables. Es decir, el periodo temporal analizado podía llegar a ser determinante.

Pero, la cuestión central del trabajo de Gombola era determinar si el CFFO era una variable importante a la hora de predecir el fracaso empresarial a partir de mediados de los años 70. Y utilizando MDA llegó a la conclusión de que el CFFO era un predictor de la insolvencia pero en el muy corto plazo, siendo este hallazgo consistente con el obtenido por Casey y Bartczak (1985).

En los estudios clásicos sobre predicción de insolvencia se empleaba MDA como técnica de clasificación. Esta técnica, para que sea efectiva, requiere de diversos supuestos entre los que

se encuentra el de normalidad de los atributos o variables empleadas y, en consecuencia, las conclusiones que se obtengan, si estos supuestos no se cumplen, se encuentran bajo sospecha.

En este sentido, Karels y Prakash (1987) se propusieron indagar en los trabajos clásicos previos con los objetivos de: (1) Averiguar si los ratios empleados en los estudios anteriores satisfacían esta condición de normalidad; (2) En caso de que no la satisficieran, se propusieron reunir un conjunto de ratios que sí las cumplieran y, en todo caso, que se desviaran lo menos posible de ese supuesto; (3) Utilizar este conjunto de ratios para comparar los resultados generados por el modelo discriminante frente a otros estudios.

La evidencia obtenida sugirió que una selección de ratios no efectuada a propósito (ad hoc) puede mejorar significativamente la predicción de insolvencia. Hay que reseñar que los ratios que terminaron siendo seleccionados fueron muy diferentes a los tradicionales (a excepción de los ratios de Fondo de Maniobra) usados por Altman y Deakin. Asimismo, dos ratios de cash flow, cash flow por acción y ventas/inventarios, que no suelen ser empleados nunca, se identificaron como variables discriminatorias significativas.

Un resultado bastante sorprendente fue la inclusión del valor de mercado de las acciones comunes con un retraso de tres años. Aquí la cuestión a dilucidar era si el mercado podía anticipar con tanta antelación la quiebra o si era una mera coincidencia de los datos. Basado en esta evidencia parece que se podía aplicar un análisis discriminante simple con una

considerable mejora en la capacidad predictiva y con menos problemas a la hora de la selección de los ratios y el muestreo.

Hing-Ling Lau (1987) también aportó innovaciones a los trabajos clásicos. El modelo que presentó desarrollaba otros modelos de predicción de insolvencia clásicos (Beaver, Altman, Ohlson, Zavgren) de dos formas: (1) Definió cinco estados financieros para describir la situación financiera de una empresa, en vez de considerar la situación dicotómica empresa fallida/empresa sana; (2) En lugar de clasificar las empresas dentro de una categoría determinada, el nuevo modelo estimaba las probabilidades de que una sociedad se situara en uno de esos cinco estados. Para ello, se definieron cinco escenarios financieros en este estudio: 0 Estabilidad Financiera; 1 Omisión del pago o reducción de los dividendos; 2 Impago o impago de préstamos; 3 Protección bajo el Capítulo X o XI de la Ley de Quiebra de Estados Unidos; 4 Bancarrota y Liquidación. Como se puede comprobar, desde los estados 1 a 4 va creciendo el nivel de las tensiones financieras.

Partiendo de este planteamiento, Hing-Ling Lau construyó tres modelos de predicción con tres horizontes diferenciados. Los modelos "Año 1" se construyeron con información financiera de 1974-1975 para predecir los problemas financieros en 1976. Igualmente los modelos "Año 2" y "Año 3" se construyeron con información de 1973-1974 y de 1972-1973, respectivamente, para predecir los problemas financieros en 1976. La muestra de validación fue usada para chequear la capacidad de estos modelos de predecir los problemas financieros en 1977.

Los modelos de predicción se construyeron con una "muestra original" y posteriormente fueron chequeados con una

muestra de validación. Cada muestra contenía 350 empresas en el estado 0, y 20, 15, 10 y 5 empresas en los estados 1, 2, 3 y 4, respectivamente.

Hay que tener en cuenta que la aplicación de un modelo probabilístico a un grupo de "n" empresas nos da "n" resultados, y que el resultado de la predicción se encuentra representado por la suma de estos "n" resultados (SSn) así como por el ratio SSn/n (dado que n es el máximo número de resultados).

Los resultados de este trabajo en la muestra de testeo mostraron diferentes porcentajes de clasificación en cada uno de los estados definidos, teniendo en cuenta el horizonte temporal elegido. Así, en el año previo a la quiebra, el porcentaje de acierto en el estado "0" fue del 96,0%, en el estado "1" un 86,65%, en el estado "2" un 88,33%, en el estado "3" del 94,5%, en el estado "4" del 34,8% y en la clasificación global de un 94,0%. En el tercer año previo a la quiebra, el porcentaje de acierto en el estado "0" fue un 95,6%, en el estado "1" un 82,6%, en el estado "2" un 82,3%, en el estado "3" un 75,2%, en el estado "4" un 56,8% y en la clasificación global un 93,55%.

Aziz et al. (1988) afirmaron que los ratios incluidos en los modelos de predicción de insolvencia se basaban más en criterios de orden práctico que en una teoría que sustentara esta elección. Este trabajo se basa en el de Gentry et al. (1985), que a su vez, se sustenta en un modelo de flujos de caja desarrollado por Helfert (1982) para obtener las variables y ratios con los que generar un modelo multivariable de predicción de insolvencia. Este estudio es similar al de Gentry et

al. (1985) pero con la diferencia de que se basa en la identidad de cash flow desarrollada por Lawson (1971). El objetivo del trabajo consistía en ver la capacidad de la identidad de Lawson para discriminar entre empresas quebradas y sanas, mediante MDAy Logit.

Para ello, usaron la base de datos de Standard and Poor´s 1983 Industrial Annual COMPUSTAT Research File para extraer las empresas que quebraron durante el periodo 1971-1982. De ella, se extrajeron 49 empresas que fueron emparejadas con otras 49 empresas sanas. La muestra era de carácter multisectorial puesto que también estaban incluidas empresas de servicios y de transporte. En total, un conjunto de 98 empresas.

Mediante MDA, las variables que presentaron un mayor poder de clasificación fueron los Impuestos y el Cash Flow Operativo. Los modelos presentaron una elevada capacidad de predicción en un rango desde el 80,9% en el quinto año hasta el 88,8% un año antes de la quiebra.

Mediante Logit los resultados obtenidos oscilaron entre el 91,8% de acierto global obtenido en el año previo a la quiebra al 80,9% de acierto global cinco años antes de la quiebra.

Se evidenció así que el modelo CFB de Aziz et al. (1988) se comparaba favorablemente con el modelo Z. El nivel de precisión global fue muy semejante, pero presentaba la ventaja diferencial de predecir hasta 5 años antes de la quiebra, por lo que era capaz de proporcionar una señal de aviso con mayor antelación. A esta mayor capacidad predictiva había que

añadirle el hecho de que la selección de variables se basó en el concepto de flujo de caja de Lawson.

Hopwood et al. (1989) examinaron la relación entre el informe de auditoría y el fallo empresarial. Partían del convencimiento de que la naturaleza de las opiniones cualificadas del auditor pueden señalar el fallo empresarial.

Este estudio, usando una aproximación log-lineal, analizó la asociación entre la quiebra y las opiniones de auditoría mediante un modelo univariante y otro modelo multivariante. Para el año inmediatamente previo a la insolvencia, el modelo univariante indicó una relación entre la salvedad “going-concern” y otras salvedades con la quiebra. Los tests de los modelos multivariantes, los cuales usaban variables de los informes de auditoría junto con otras, indicaron igualmente una relación de la quiebra tanto con la calificación “going-concern” como con otro tipo de salvedades. Cuando las variables de los informes de auditoría fueron testadas con otras variables específicas de un modelo de predicción de insolvencia, las salvedades y la calificación “going-concern” siguieron estando relacionadas con el fallo empresarial.

Igualmente, los análisis efectuados consideraron la capacidad predictiva de los modelos, testando la sensibilidad de los resultados con respecto a los errores tipo I y tipo II. Los resultados indicaron que, en la mayoría de los casos, el modelo que sólo considera la opinión del auditor era la menos onerosa de las alternativas en un periodo de hasta tres años.

Aziz y Lawson (1989) propusieron un modelo de cash flow similar al de Gentry et al. (1985) en el sentido de que ambos

partían de una definición de cash flow. Este estudio testeó la validez de la identidad de Lawson, incluyendo flujos de capital externos. Una diferencia crítica entre este estudio y los de sus predecesores es su enfoque en la formulación de hipótesis. Se compararon los resultados de cuatro modelos (Z, Zeta, CFB y mixtos) y sus diferencias fueron testadas usando las técnicas estadísticas habituales.

Este trabajo aclaró que el cash flow operativo, el cash flow de la financiación, el cash flow de la inversión y los impuestos son variables importantes a la hora de predecir la insolvencia. En este sentido, la sugerencia de Casey y Barack (1985) de no enfatizar la importancia del cash flow a la hora de predecir la salud financiera de las empresas estaría en completo desacuerdo con este enfoque.

Pioneros en la aplicación de NN a la predicción de insolvencia, Odom y Sharda (1990) se propusieron llevar a cabo una comparación de los resultados obtenidos mediante NN y métodos más tradicionales como MDA. La predicción se llevó a cabo para 1 año antes de la quiebra.

La muestra estaba compuesta por 129 empresas, 65 de las cuales entraron en bancarrota en el periodo 1975-1982 y 64 restantes, sanas. La muestra se obtuvo de la base de datos del Moody's Industrial Manuals. Se hicieron dos submuestras a partir de esta muestra principal de 129 empresas: Una muestra de entrenamiento compuesta por 74 empresas (38 quebradas y 36 sanas) y una muestra de testeo de 55 empresas (27 quebradas y 28 no quebradas), emparejándolas por industria y año.

Para garantizar la robustez de los resultados, la proporción muestra de entrenamiento/muestra de testeo se fue alterando para reflejar mejor la realidad con respecto a la proporción empresas sanas/empresas quebradas. Se crearon tres grupos, en función de la proporción, 50/50, 80/20, 90/10. Los resultados revelaron que las NN mostraron resultados más robustos en muestras más pequeñas, obteniendo mejores resultados en la muestra de testeo.

Nuevamente, dentro del ámbito de los estudios que relacionaban la opinión de auditoría con la predicción de insolvencia, Koh y Killough (1990) se propusieron como objetivo en este estudio tratar las dificultades que tiene para los auditores otorgar la calificación de "going-concern" a las compañías y que posibles medidas podrían ayudar en esta tarea. En concreto, construyeron un modelo de predicción de insolvencia mediante MDA para ver las posibilidades y limitaciones de esta herramienta en el trabajo de los auditores.

La muestra se compuso por 70 empresas, la mitad de las cuales eran empresas en quiebra (Wall Street Journal Index) durante el periodo 1980-1985. El emparejamiento se llevó a cabo en función del tamaño de los activos y del sector (código SIC), siendo los datos financieros empleados los de un año antes de la quiebra. Para comprobar el grado adecuación de la función discriminante para detectar problemas de empresas (a la hora de predecir la quiebra) a la valoración "going-concern" de los auditores, se extrajo una nueva muestra de 400 compañías a partir de COMPUSTAT durante el periodo 1980-1985, donde, de estas 400 compañías, 14 de ellas estaban quebradas.

Así, la función discriminante incorporó cuatro variables: Quick ratio, Beneficios Retenidos/Total Activo, Beneficio neto por acción y Dividendos por acción. Los resultados de la clasificación arrojaron un error global del 7,35%, un error tipo I del 8,82% y un error tipo II del 5,88%.

La comparación del modelo MDA con la valoración "going-concern" de los auditores mostró que el porcentaje de acierto era similar, tanto en la predicción del modelo como en la opinión de los auditores por lo que se refiere a las empresas sanas (88,60% frente a 88,86%). Sin embargo el modelo MDA detecta mucho mejor a las empresas que quiebran (78,57% vs. 21,43%).

Los resultados obtenidos podrían explicarse debido a los elevados costes que, para las compañías auditoras, tiene el equivocarse en cuanto a la continuidad de una empresa. En este sentido, los autores sugieren que sería importante disponer de una herramienta de predicción de la insolvencia específica para ayudar al auditor en su tarea.

Hay que recalcar asimismo que la calificación "going-concern" no implica la entrada en quiebra de una sociedad y, sin embargo, presenta riesgos reputacionales y de imagen importantes para la compañía y para el auditor en caso de no formular esta valoración, debiendo hacerlo. Además, hay que apuntar que el buen juicio del profesional es muy difícil de sustituir.

Gilbert et al. (1990) pretenden distinguir entre las empresas que se encuentran estresadas financieramente y terminan en quiebra de las que se salvan finalmente de

quebrar. Parten de la idea de que un modelo de predicción de insolvencia desarrollado usando una muestra de empresas en quiebra y de empresas tomadas al azar es incapaz de distinguir a las empresas estresadas financieramente de las que finalmente quiebran. Esto demostraría que las empresas que entran en quiebra lo hacen, además de por razones financieras, por otras de carácter no financiero. Asimismo, las variables financieras que no son capaces de diferenciar entre empresas quebradas y financieramente estresadas son diferentes de las que son capaces de diferenciar entre empresas en quiebra y empresas sanas escogidas al azar. Por otro lado, los autores mostraron cómo las variables de cash flow aportaban capacidad explicativa a los modelos, al contrario de lo que se venía apuntando por otros autores.

Para abordar todas estas cuestiones seleccionaron tres grupos de empresas: a) un grupo de empresas en quiebra; b) un grupo de empresas escogido aleatoriamente; c) un grupo de empresas financieramente débiles, pero no en quiebra.

Los tres grupos de empresas se utilizaron para formar, a su vez, dos tipos diferentes de muestras que fueron usadas posteriormente para el análisis. Las 76 empresas quebradas y las 304 empresas que se obtuvieron aleatoriamente formaron la muestra denominada "Quebradas/Aleatorias". Las 76 empresas quebradas se combinaron con las 304 empresas estresadas financieramente para crear la muestra "Quebradas/Estresadas". Cada una de estas muestras se dividió, a su vez, en una muestra de entrenamiento y en una muestra de testeo. En el caso de la muestra "Quebradas/Aleatorias" la muestra de entrenamiento se formó por 52 quebradas y 208 aleatorias. La muestra de testeo se compuso de las 24 quebradas y 96

aleatorias restantes. Lo mismo se hizo para la muestra "Quebradas/Estresadas".

Así, los resultados del modelo Logit para la muestra de empresas "Quebradas/Aleatorias" incluyó tres variables como significativas (EBIT/Total Activo, Caja de las operaciones/Total Pasivo, Fondos Propios/Total Pasivo). Por su parte, la tasa de acierto fue del 88,5% en la muestra de entrenamiento y del 90,8% en la muestra de testeo. Asimismo, el error tipo I fue del 37,5% en la muestra de testeo (32,7% en la muestra de entrenamiento). Sin embargo, cuando este modelo se aplicó a la muestra "Quebradas/Estresadas" los resultados empeoraron notablemente, cayendo la tasa global de acierto, con un porcentaje de error global del 33,3%, un error tipo I del 37,5% y un error tipo II del 32,3%.

Por lo que se refiere a los resultados del modelo Logit para la muestra "Quebradas/Estresadas", el modelo incluyó estas variables como significativas: Caja de las operaciones/Pasivo corriente, Caja/Total Activo, Fondos Propios/Total Pasivo y Beneficios Retenidos/Total Activo. La tasa de acierto fue del 18,1% en la muestra de entrenamiento y del 21,7% en la muestra de testeo. El error tipo I fue 69,2% en la muestra de testeo (70,8% en la muestra de entrenamiento).

Como señala Taffler (1984), los resultados de los modelos de predicción de insolvencia deben de ser interpretados como descripciones de estrés financiero más que como predicciones de insolvencia per se. Sin embargo, los resultados de este estudio sugieren que si el objetivo es identificar probables quiebras dentro de un grupo de compañías con problemas, los resultados de estos modelos resultan francamente muy pobres.

Otro de los trabajos pioneros en la introducción de las NN fue el de Coats y Fant (1992). El objetivo era describir lo que eran las NN y su aplicación al campo de la predicción de la insolvencia empresarial.

Para valorar la potencia clasificatoria de las NN se tomó una muestra de empresas que habían sido identificadas, por sus auditores, como financieramente estresadas. Se utilizaron datos financieros de los tres años previos a entrar en esa situación de estrés financiero y se tomaron otras 94 empresas sanas. Después de esta selección, ambas muestras (empresas sanas y empresas en dificultades) se dividieron en dos grupos. Cada uno de los grupos estaba formado por 47 empresas financieramente estresadas y 47 empresas sanas. Un grupo sirvió de muestra de entrenamiento y el otro como muestra de testeo.

Las NN fueron capaces de predecir correctamente el 91% de las empresas en dificultades financieras y el 96% de las empresas sanas. Por otro lado, MDA fue capaz tan sólo de acertar el 72% de las empresas estresadas y el 89% de las empresas sanas. Si atendemos a los errores, NN presentó un error tipo I del 9% frente al 28% de MDA. Además, NN presentaron un error tipo II del 4% frente al 11% de MDA.

Fletcher y Goss (1993) ilustraron el desarrollo de un modelo de predicción de insolvencia empresarial usando una clase particular de NN como es la Red de Propagación hacia atrás o de Retropropagación. Los datos para la elaboración de este trabajo fueron extraídos de un estudio anterior de Gentry. Se seleccionaron 18 empresas fallidas junto con 18 empresas

sanas y se emparejaron de acuerdo con el tamaño, según las ventas y volumen de activo.

Los cinco modelos NN obtenidos predijeron todos mejor que el modelo Logit con el que se comparaba. En concreto, el modelo de NN que mayor nivel de predicción alcanzó obtuvo una tasa de acierto del 82,4% frente al modelo Logit que arrojó un resultado del 71,3%.

Para Theodossiou (1993), las señales de que una empresa se deteriora financieramente se desarrollan en el tiempo. Normalmente, el cambio requiere de varios años, por lo que consideró necesario desarrollar modelos secuenciales que hicieran posible detectar tales cambios tan pronto como fuera posible. En este trabajo Theodossiou presentó un procedimiento secuencial para detectar un cambio en la media de una serie temporal multivariante, y muestra cómo este procedimiento puede ser usado para predecir la tendencia de una empresa hacia la quiebra.

El procedimiento está basado en el trabajo previo de Healy (1988) de sumas acumulativas (CUSUM) y Shumway (1988) sobre un procedimiento de series temporales multivariante. Presentó un modelo secuencial (dinámico) basado en el análisis CUSUM para la predicción de las empresas con problemas financieros. Específicamente, este modelo podía ser usado para detectar el punto a partir del cual las variables financieras de una empresa pasaban de un "buen comportamiento" a un "mal comportamiento". El modelo puede ser visto como una extensión dinámica de MDA.

Adicionalmente, el modelo presentaba como característica deseable el tener una memoria corta para el buen comportamiento de las empresas a lo largo de los años y una memoria larga para el mal comportamiento. En consecuencia, alertaba de manera rápida al analista para que pudiera investigar y valorar la empresa de manera más cercana y minuciosa. Cuando se compara el modelo CUSUM con MDA, el primero se reveló como claramente superior.

En el trabajo de Johnsen y Melicher (1994) se continuó la línea de investigación que trata de trascender la tipología clásica de estudios sobre predicción de insolvencia, superando la tradicional clasificación dicotómica. En este sentido, la quiebra, después de todo, no sería más que un caso extremo de dificultad financiera. Por ello, los problemas financieros se describen mejor como un estado continuo que se sitúa entre lo "financieramente débil" hasta la "quiebra", con la posibilidad de varios estados intermedios de "debilidad financiera".

En esta línea, Lau (1987) fue el primero en superar la dicotomía empresa sana/empresa quebrada y definió 4 tipos de estados de debilidad financiera. Hay que tener presente que este tipo de modelos presentan, a su vez, más tipos de errores de clasificación.

El trabajo de Johnsen y Melicher (1994) pretendió replicar los estudios previos en esta línea de investigación utilizando el mismo tipo de muestreo y construyendo un modelo de predicción de insolvencia compuesto de dos estados. Se compararon los resultados con los de un modelo multiestado con una categoría añadida explícita de "financieramente débil" que es incluida en la muestra. El segundo objetivo del trabajo

fue determinar el valor añadido que suponía la clasificación de información secundaria en la predicción de los cambios en la condición financiera.

La muestra utilizada para este estudio estaba compuesta por tres grupos de empresas: Empresas quebradas (112 empresas; 17% del total); empresas sanas (293 empresas; 44,4% del total) y empresas financieramente débiles (255 empresas; 38,6% del total). La selección de las empresas se hizo para el periodo 1970-1983 y se tomaron los datos financieros de los últimos cinco años previos a la quiebra.

El estudio aportó tres conclusiones muy claras: 1) Los errores de clasificación se pueden reducir añadiendo un estado adicional a la quiebra; 2) También es importante destacar que los tres estados de salud financiera parecen ser "independientes" y 3) Se extrajo una importante información sobre las probabilidades de mejorar o empeorar la condición financiera basada en la situación de partida.

Altman et al. (1994) analizaron la comparación entre las metodologías estadísticas clásicas para la clasificación y predicción de la quiebra empresarial (MDA y Logit) con las redes neuronales (NN). Los resultados mostraron un grado bastante equilibrado de precisión y otras características beneficiosas entre MDA y NN. Ambos tipos de análisis mostraron más de un 90% de acierto en la clasificación en la muestra de testeo, concluyendo que debería seguirse esta línea de investigación usando estas dos técnicas y sugiriendo una aproximación combinada para un refuerzo predictivo.

En el marco del proyecto de la Central de Balances del Banco de Italia, se tomaron dos muestras para calcular sendas funciones discriminantes. La primera función discriminante pretendía clasificar empresas sanas y quebradas y la segunda, empresas vulnerables y quebradas. Se tomaron, como muestra de entrenamiento, 404 empresas industriales sanas y 404 empresas industriales quebradas. A su vez, para la muestra de testeo se tomaron 150 empresas sanas y otras tantas quebradas.

Los resultados obtenidos en las dos funciones mediante el análisis discriminante arrojaron unos porcentajes de acierto, para la muestra de testeo, que se situaron, para la primera función, en el 95,1% un año antes de la quiebra y, para la segunda función, en el 81% un año antes del fallo. Mediante NN se obtuvo un porcentaje de acierto en la clasificación del 95,3% un año antes de la quiebra.

Probablemente, la conclusión más importante del trabajo es que NN no son una técnica claramente dominante sobre los métodos tradicionales, como MDA. Los autores especularon que una aproximación más equilibrada sería más adecuada teniendo en cuenta los problemas de "caja negra" que presentan NN.

Otro estudio que realizó un análisis comparativo de la capacidad predictiva de NN frente a MDA en la predicción de la insolvencia fue el de Wilson y Sharda (1994). La muestra que emplearon estaba formada por 129 empresas (65 quebradas y 64 sanas) durante el periodo 1975-1982, pertenecientes al sector industrial en sentido amplio. Los datos utilizados de las empresas quebradas fueron obtenidos del Moody's Industrial

Manual, en concreto, los estados financieros un año antes de entrar en quiebra.

Los resultados obtenidos mediante NN se mostraron netamente superiores al análisis discriminante, con porcentajes que oscilaban entre el 97,5% y el 72,62% en el caso de NN, frente al 91,59% y 72,0% en el caso de MDA.

Hay que reseñar que se llevaron a cabo diferentes composiciones de las muestras, variando la proporción de componentes de la muestra de entrenamiento y de la muestra de testeo. También que la composición de la muestra de entrenamiento tuvo efectos determinantes en la capacidad predictiva de NN.

Por su parte, el objetivo del estudio de Ward (1994) fue determinar por qué la variable NOF (Net income+Depreciation+Amortization/Total debt) es un predictor tan potente del estrés financiero.

En el trabajo se discutieron dos posibles explicaciones: (1) La primera es que era probable que la variable NOF pudiera ser una medida simple de cash flow operativo, y (2) Es posible que sea un predictor poderoso de estrés financiero debido a que sea una mejor medida de ingreso económico que el beneficio después de impuestos.

Este trabajo también amplió la investigación previa sobre estrés financiero mediante el desarrollo de modelos de predicción de cuatro estados ordinales de estrés financiero, en lugar de emplear los tradicionales modelos dicotómicos.

Los resultados mostraron que los modelos ordinales de estrés financiero ofrecían una respuesta viable a la hora de enfocar de una forma más realista el problema de la insolvencia empresarial.

En el estudio de Boritz y Kennedy (1995) se examinó la efectividad de diferentes tipos de NN a la hora de predecir la insolvencia empresarial.

Se consideraron dos aproximaciones para entrenar NN: la retropropagación y la teoría de la estimación óptima. Dentro del mecanismo de retropropagación se testearon cuatro modelos (Retropropagación, Retropropagación de Nexo Funcional con senos, Retropropagación Reducida y Retropropagación Predictiva Acumulativa). Y NN se compararon con los métodos tradicionales de predicción como MDA, Logit y Probit.

Los resultados mostraron que los errores tipo I y tipo II variaban considerablemente según fueran las técnicas empleadas. Así, NN presentó el menor porcentaje de error tipo I y el mayor porcentaje de error tipo II, en tanto que las técnicas tradicionales mostraron el patrón contrario.

Los modelos NN con algoritmo de retropropagación obtuvieron unos porcentajes intermedios de error tipo I y error tipo II, demostrándose que el resultado de estos modelos era muy sensible a la elección de las variables y que no se podía confiar en que los modelos eligieran las variables más importantes. De igual forma, los resultados fueron muy sensibles a los errores de muestreo.

En los Estados Unidos los auditores están obligados a modificar sus informes si el periodo de existencia de las empresas se estima inferior al año (salvedad "going-concern"). Pero la predicción del estado de "going-concern" para un auditor puede ser un asunto muy delicado. Incluso cuando los auditores están en condiciones de predecir problemas potenciales de quiebra, suelen ser habitualmente reticentes a revelar esta información debido a las posibles consecuencias de una salvedad de ese tipo. De hecho, las investigaciones que habían examinado la situación de las compañías durante el año posterior a haber recibido una salvedad de "going-concern" por parte de sus auditores reveló que aproximadamente entre el 25% y el 38% de las compañías examinadas acaban quebrando un año después de haber recibido una salvedad del tipo "going-concern".

Siguiendo esta línea de investigación, McKee (1995) trató de desarrollar un modelo muy preciso y, a la vez, teóricamente consistente de predicción de la insolvencia. Para ello, empleó un algoritmo inductivo diseñado para extraer el aprendizaje de casos del mundo real. La investigación señaló que los auditores experimentados sólo pueden predecir la insolvencia correctamente en el 83% de los casos. Asimismo, indicó que los auditores solo informan de las posibles quiebras en el 45% aproximado de los casos, concluyendo que no es fácil para el auditor predecir e informar sobre las probabilidades de quiebra de las compañías. El trabajo desarrolló un modelo simple basado en dos únicos ratios (calculados sobre la información disponible) que proporcionó una precisión en la predicción del 97% en 60 compañías privadas.

En el trabajo de Leshno y Spector (1996) se compararon las capacidades predictivas de varios modelos NN. Los modelos examinados diferían en la amplitud de los datos, la técnica de aprendizaje y el número de iteraciones. Asimismo, las capacidades de cada modelo fueron también comparadas usando el clásico MDA.

Dado que el número de empresas quebradas es limitado, se tomaron estados financieros de varios periodos precedentes al evento de quiebra. Si bien los estados financieros del periodo de quiebra suelen proporcionar más información, incorporar los estados financieros de distintos periodos puede mejorar los resultados. Las capacidades predictivas de los modelos se ven realizadas mediante el uso de técnicas de aprendizaje mejoradas. Sin embargo, si la técnica de aprendizaje es demasiado "fuerte" el modelo se hace muy específico "con respecto a la muestra" y, por tanto, pierde capacidad de generalización.

Finalmente, el modelo NN preferido es el que usa una expresión funcional de segundo orden junto con una activación conjunta de segundo orden y 100 iteraciones de aprendizaje. Este modelo demostró ser más preciso que MDA.

En su trabajo, Serrano-Cinca (1996) aplicó un tipo específico de NN llamado "Mapa Característico Auto-Organizado" para el análisis de la información financiera. Este modelo lleva a cabo una proyección de un espacio multidimensional sobre el plano, de tal modo que los puntos que se encuentran unos cerca de otros se corresponden con patrones similares en los inputs de información. El uso de NN se

complementó, en este caso, con un estudio estadístico de la información disponible.

Sobre este planteamiento base, se desarrolló un Sistema de Apoyo a la Decisión (Decision Support System, DSS) para el análisis de los estados financieros que incluyó un Análisis Lineal Discriminante y un Perceptrón Multicapas (MLP) para delimitar las regiones de insolvencia. El DSS propuesto mejoraba el Z análisis tradicional y proporcionó una visión intuitiva y gráfica sobre el riesgo de insolvencia, las características financieras de la empresa y la tipología de empresa a la que pertenecía.

Chang et al. (1996) desarrollaron tres modelos híbridos de NN para la predicción de la insolvencia. Los modelos propuestos fueron: (1) Una NN asistida por MDA; (2) Una NN asistida ID3 y (3) Un SOFM (una NN característica auto-organizada). Tanto el modelo (1) como el (2) son modelos NN que operan con las variables input seleccionadas por los métodos MDA e ID3, respectivamente. El SOFM combina un modelo de retropropagación (modelo supervisado) con un modelo SOFM (con aprendizaje, sin supervisar). Los resultados del modelo NN híbrido son evaluados usando MDA e ID3 como referencia.

Los resultados empíricos, usando los datos de quiebras de Corea del Sur, mostraron que los modelos NN híbridos eran modelos muy prometedores, tanto en términos de la precisión de sus resultados como de su adaptabilidad. Asimismo, señalaron que la capacidad de predecir los problemas financieros podía ser reforzada mediante el uso de modelos que tuvieran en cuenta la correlación en los datos, incorporasen

información de más de un periodo e incluyeran variables explicativas de carácter estacionario.

El artículo de Kahya y Theodossiou (1999) desarrolló un modelo estacionario de estrés financiero para empresas manufactureras y comerciales procedentes del AMEX y NYSE, basado en la metodología de series temporales de sumas acumulativas (CUSUM). El modelo poseía la particularidad de diferenciar entre cambios en las variables financieras de una empresa, que son el resultado de la correlación serial, y los cambios, que son el resultado de cambios permanentes en la estructura de las medias de las variables como consecuencia de los problemas financieros. Los tests llevados a cabo mostraron que el modelo era robusto en el tiempo y que mejoraba los resultados de modelos similares basados en MDA y Logit.

El modelo CUSUM podía ser visto como una versión dinámica de MDA. Así, una característica muy positiva del modelo CUSUM es que tiene una memoria muy corta con respecto a los buenos resultados de la empresa a lo largo de los años y, en cambio, una memoria muy larga con respecto a los malos resultados. Esta característica de memoria hace que el modelo sea muy sensible a los cambios negativos en las condiciones financieras. En consecuencia, alerta muy rápidamente al analista financiero, quien puede llevar a cabo una investigación mucho más minuciosa y una valoración más realista de su situación.

El trabajo de Charalambous et al. (2000) compara el resultado predictivo de tres tipos de NN con el rendimiento mostrado por Logit y el algoritmo de retropropagación. Los resultados mostraron que la capacidad predictiva de los

métodos NN superaba a los acreditados por las técnicas clásicas. Para este estudio, la muestra se compuso por 139 empresas sanas y 139 empresas quebradas durante el periodo 1983-1994, llevándose a cabo el emparejamiento atendiendo a criterios de tamaño, año de entrada en quiebra y sector industrial considerado. Las muestras se fragmentaron, a su vez, en muestra de entrenamiento y muestra de testeo. La muestra de entrenamiento estaba compuesta por todas las empresas, tanto quebradas como solventes, incluidas en la muestra global durante el periodo 1983-1991 (un total de 192 empresas), en tanto que la muestra de testeo consistió en las 86 empresas restantes del periodo 1992-1994.

Los resultados mostraron que las tres tipologías de NN aplicadas en este estudio proporcionaban unos resultados superiores a los aportados por Logit y por el algoritmo de retropropagación.

Nasir et al. (2000) también pretendieron comprobar la superioridad de NN sobre las técnicas tradicionales de predicción de la insolvencia. La muestra se compuso de 15.000 empresas quebradas del Reino Unido, empleando datos financieros completos durante el periodo 1992-1996. Dichas empresas se emparejaron con igual número de empresas sanas para, a continuación, ser ordenadas según tamaño y subsector industrial.

El artículo de Shumway (2001) propuso un modelo de riesgo (Hazard Model) en contraposición a los modelos estáticos tradicionalmente utilizados por la literatura. Los modelos estáticos tan sólo pueden considerar un conjunto de variables explicativas para cada empresa. Además, la mayoría de los

investigadores eligen el momento de la observación, que suele coincidir con un año antes de la quiebra, lo que introduce un sesgo en la estimación que es lo que el modelo de riesgo propuesto pretende evitar.

Para ello, compuso una muestra formada por empresas cotizadas durante el periodo 1962-1992, principalmente del sector industrial. Se tomaron como empresas quebradas todas las que fueron dadas de baja de NYSE o AMEX y quebraron en los 5 años siguientes. La muestra final contenía 300 empresas que quebraron en el periodo 1962-1992.

Los resultados de la comparación con el modelo de Altman mostraron que su modelo era más preciso ya que clasificaba casi el 70% de todas las quiebras en el decil de probabilidad más alto. Asimismo, clasificó el 96,6% de las empresas quebradas por encima de la probabilidad mediana.

Finalmente, en la comparación de los modelos usando variables contables y de mercado se obtuvieron los mejores resultados. Este modelo clasificó correctamente el 75% de las empresas quebradas en el decil más alto y sólo clasificó el 3,5% de las empresas quebradas por debajo de la probabilidad mediana.

Yi-Chung et al. (2005) usaron dos técnicas de inteligencia artificial, MLP y la Red de Función de Base Radial (RBFN) para resolver la clasificación y predicción de la quiebra en una muestra de empresas inglesas. El resultado mostró que la RBFN era superior a MLP, a MDA y a Probit. Para ello, se seleccionaron inicialmente 904 empresas inglesas del sector industrial. De éstas, 353 empresas quebraron o fueron adquiridas por otra

compañía entre marzo de 1985 y marzo de 1994. Los datos de las empresas quebradas fueron tomados de los estados financieros un año antes de la quiebra y, tras la aplicación de determinados filtros, se seleccionaron 32 empresas quebradas como muestra definitiva. A continuación, se escogieron 45 empresas sanas sobre la base de disponer de 9 años de datos financieros ininterrumpidos.

RBFN se mostró como un método superior a otros. Esto podría ilustrar que la RBFN podría emplearse para soslayar los problemas de ajustes de curva. Los resultados analíticos mostraron que era posible su empleo potencial considerando el carácter no aditivo de la interacción entre atributos, como ocurre en los sistemas de reglas fuzzy en la predicción de insolvencia.

En el estudio de Pindado et al. (2008) se desarrolló un modelo ex ante para la estimación de los problemas financieros, presentando a su vez una definición de dificultad financiera, independiente de sus consecuencias legales, junto con un modelo soportado teóricamente para la predicción de los problemas financieros, además de una metodología que usaba datos de panel para eliminar la heterogeneidad no observable.

Este estudio ofreció una aproximación a la hora de estimar la probabilidad de tener problemas financieros que podía ser aplicado a diferentes países y periodos temporales. El modelo se mostró muy estable en cuanto a la significatividad, a la estabilidad de los coeficientes y al poder de clasificación para diferentes periodos y países. Sin duda, el desarrollo de una medida robusta y estable de la probabilidad de tener problemas financieros fue una aportación significativa a la literatura en el

sentido de que tiene múltiples aplicaciones en muchas áreas de las finanzas.

Muller et al. (2009) compararon los resultados de varias técnicas de predicción de insolvencia. Las técnicas examinadas fueron MDA, la partición recursiva (RP), Logit y NN. Introdujeron, como concepto novedoso, el "Coste Normalizado de la Quiebra" (NCF), teniendo en cuenta que el coste del error tipo I se suele situar en el entorno de 25-35 veces el error tipo II.

Los datos y la selección de variables se basaron en el estudio de Steyn-Bruwer y Hamman (2006). La muestra se compuso por las empresas sudafricanas cotizadas de todos los sectores excluyendo la minería, el sector financiero y el inmobiliario, siendo el periodo considerado el comprendido entre 1997 y 2002. Adicionalmente, se utilizó el periodo 2003-2006 para comprobar el porcentaje de empresas quebradas acertado.

Los resultados, mediante MDA, mostraron que el NCF fue inferior cuatro años antes de la quiebra, representando algo menos del doble que el registrado un año antes. Mediante Logit, el NCF fue inferior cuatro años antes de la quiebra, representando igualmente algo menos del doble que el registrado un año antes. Mediante RP, el NCF fue inferior cuatro años antes de la quiebra, representando un 65% menos que el registrado un año antes. Y finalmente, mediante NN, el NCF fue inferior cuatro años antes de la quiebra, representando un 62,2% menos que el registrado un año antes.

Los resultados concluyeron que MDA y RP fueron los métodos que predijeron el mayor porcentaje de empresas quebradas y que, por tanto, presentaron el menor NCF. En cambio, Logit y NN produjeron el mejor porcentaje global de acierto. Cuando los datos se referían a uno, dos, tres y cuatro años antes de la quiebra, en vez de estar referidos al ciclo económico, la calidad de la predicción mejoró.

Shuk-Wern et al. (2011) desarrolló un modelo que pudiera predecir el estrés financiero entre las empresas públicas cotizadas en Malasia mediante Logit. Los resultados mostraron que 5 ratios financieros resultaron ser significativos y útiles para predecir el estrés financiero en Malasia. La precisión global del modelo alcanzó el 91,5%, lo que demostraba que Logit era una herramienta muy fiable para la predicción del estrés financiero.

La muestra para este estudio se extrajo de una selección de empresas cotizadas en la Bolsa de Malasia y con evidencias de estrés financiero desde 2001 a 2007. La muestra estaba compuesta por 105 empresas de diversos sectores (productos industriales, comercio y servicios, inmobiliarios, productos de consumo, construcción y finanzas) y se subdividió a su vez en dos submuestras: una muestra de entrenamiento de 100 compañías para llevar a cabo el modelo (periodo 2001-2006) y una muestra de validación de 5 compañías (2007). El emparejamiento se llevó a cabo al 50% teniendo en cuenta el subsector y el tamaño.

Los resultados mostraron que, en la muestra de entrenamiento, el modelo registró un error global del 8,5%, presentando un error tipo I del 10% y del tipo II del 7%. En la muestra de testeo, el modelo ofreció un 90% de acierto.

Los resultados obtenidos mostraron, además, que Logit resultó ser una técnica más fiable que MDA a la hora de predecir la insolvencia. Los resultados revelaron asimismo que cinco ratios: a) Rotación de activos corrientes; b) Rotación de activos totales; c) Días cobro a clientes; d) Ratio cash flow/deuda; e) Total pasivo/Total activo, resultaron ser significativos.

La predicción del riesgo de impago es importante tanto para las empresas como para los bancos en su necesidad de operar exitosamente. Chen et al. (2011) propusieron la metodología SVM para predecir la insolvencia en las empresas alemanas. Los objetivos de este trabajo fueron: (1) Investigar cuál de los ratios contables empleados eran significativos y tenían un carácter predictivo en relación con la insolvencia; (2) Demostrar que un modelo SVM bien especificado superaba a los métodos tradicionales MDA y Logit.

Algunos de los ratios empleados para estimar el modelo SVM eran novedosos por cuanto presentaban una elevada dependencia no lineal sobre el riesgo de impago, pues una dependencia lineal débil que no podía ser capturada por los modelos tradicionales como MDA o Logit.

Los datos usados en este trabajo provenían de la Base de Datos Credit Reform Database. Contenía una muestra aleatoria de 20.000 empresas solventes y 1.000 empresas insolventes durante el periodo comprendido desde 1996 a 2002, si bien los datos se encontraban concentrados en 2001 y 2002, con un 50% de las observaciones en este periodo. Los datos de las empresas insolventes fueron recopilados dos años antes de la insolvencia. Tras aplicar una serie de filtros, la muestra quedó

compuesta por 783 empresas insolventes y 9.583 empresas solventes, emparejándose de acuerdo con el sector y el tamaño.

Para el caso de un modelo univariante, los resultados mostraron que el modelo Logit presentó mayores errores tipo I, tipo II y error total que el modelo SVM. El ratio Proveedores/Compras fue el mejor predictor encontrado.

Para el caso de un modelo multivariante, el modelo SVM volvió a superar al modelo Logit. En este caso, los mejores predictores en el modelo multivariable fueron Ingresos/Total Activo, Caja/Total Activo, Variación de Existencias/Existencias, Existencias/Ventas, EBIT/Total Activo y Beneficio Neto/Ventas.

El estudio concluyó que las empresas alemanas tenían una mayor probabilidad de quiebra cuando presentaban una elevada rotación de personal, bajos beneficios, poca caja, alto endeudamiento y una elevada rotación en sus inventarios. Entre los 28 ratios manejados, el ratio de pago a proveedores fue seleccionado por el modelo SVM como el mejor predictor del riesgo de insolvencia, lo que aumentaba diferencialmente la capacidad de predicción y clasificación sobre el modelo Logit en más de un 10%.

El principal propósito del trabajo de Kwak et al. (2012) fue evaluar la capacidad de las técnicas de minería de datos. Para ello, propuso una técnica de programación lineal multicriterio usando datos de empresas coreanas en quiebra justo después de la crisis de 1997. La muestra consistió en 65 empresas quebradas y 130 empresas sanas del sector industrial durante los años 1997 a 2003.

Los resultados obtenidos sobre el modelo de Altman presentaron una tasa de error global del 12,8%. La aplicación de su metodología al modelo de Ohlson registró un error global del 16,4%. Y, finalmente, la aplicación sobre un modelo combinado Altman-Ohlson dio un porcentaje de error global del 11,3%. En definitiva, los resultados confirmaron que este método funcionaba tan bien como los métodos tradicionales, ya fuera MDA o Logit, siendo los resultados comparables a los obtenidos con técnicas como SVM o DT, si bien los resultados no eran generalizables debido a que los datos procedían de empresas en un periodo temporal peculiar.

El artículo de Sangjae y Wu (2013) presentó una investigación multisectorial de la probabilidad de insolvencia en empresas coreanas utilizando una tipología de red neuronal BNN (Back-propagation neural network). Los sectores comprendían desde el sector de la construcción, comercio minorista hasta la industria manufacturera. La precisión predictiva de la red BNN fue comparada con la que proporcionó el análisis MDA.

La muestra estaba formada por las empresas que quebraron durante el periodo 2000-2009 y que dejaron de cotizar en la bolsa de valores coreana. El resto, es decir, las empresas sanas, fueron seleccionadas de entre las compañías cotizadas y pertenecientes al mismo sector de actividad y con un tamaño similar de activos. El número total de empresas ascendió a 229, incluyendo 91 compañías quebradas. Por sectores, la muestra de estimación se compuso de la siguiente forma: construcción (24 quebradas y 48 sanas), comercio (27 quebradas y 37 sanas), industria (37 quebradas y 49 sanas). Para la muestra de validación, la composición quedó:

construcción (1 quebrada y 2 sanas), comercio (1 quebrada y 2 sanas), industria (1 quebrada y 1 sana).

Los resultados revelaron una precisión en la muestra global mediante BNN del 81,43%, siendo en la construcción de un 92%, en comercio de un 89,28% y en industria un 90,79%. Mediante la aplicación de MDA, la precisión global alcanzó un 74,82%, siendo en la construcción de un 82,01%, en comercio del 80,96% y en industria del 81,58%.

1.4.2 MODELOS CENTRADOS

En este apartado se examinarán aquellos estudios sobre predicción de insolvencia que específicamente se han basado en la predicción del fracaso empresarial en un sector económico específico, en nuestra terminología, modelos centrados.

A tal efecto, en este estudio se han considerado cinco sectores económicos: sector agricultura, sector industria, sector construcción, sector comercio y servicios, y sector hostelería.

1.4.2.1 Sector Agricultura

Dietrich et al. (2005) desarrollaron un modelo de predicción de insolvencia específico para las cooperativas agrarias. El objetivo del trabajo era comprobar si un modelo desarrollado con datos de empresas agrarias mejoraba la predicción de la insolvencia frente al resultado obtenido a partir de una muestra de compañías con inversores privados. El resultado del estudio evidenció que, efectivamente, se producía una mejora en la calidad de la predicción de la quiebra cuando

el modelo se aplicó a cooperativas agrícolas de New Brunswick (Canadá).

La metodología empleada en su estudio fue MDA, que aplicaron a una muestra obtenida del Departamento de Justicia de la Provincia de New Brunswick (Canadá) para los años 1999, 2000 y 2001. La muestra contenía las cuentas de resultados y los balances de sesenta y siete cooperativas agrícolas. En total se utilizaron trece variables para la elaboración del modelo. Doce de ellas procedían de los dos modelos predictores de quiebras más empleados en Estados Unidos y Canadá (Springate et al., 1978 Fulmer et al., 1984). Adicionalmente, emplearon la variable Fondos Propios/Total Pasivo, utilizada por Altman, que no había sido usada en ninguno de los modelos mencionados.

La evaluación de los resultados obtenidos se llevó a cabo mediante la especificación de un indicador denominado AR, que definieron como el porcentaje de predicciones acertadas sobre el número total de estados financieros testados. De acuerdo con este indicador, juzgaron la precisión de los dos modelos estimados: un primer modelo amplio con trece variables (C13) y un segundo modelo, más reducido, con siete variables (C7).

Por lo que se refiere a la capacidad de predicción de los modelos, tanto (C13) como (C7) resultaron superiores a los modelos de Springate y Fulmer. Así, el indicador AR fue del 77,63% para el modelo (C13) y del 76,13% para el modelo (C7), frente a un 70,15% que arrojó el modelo de Springate y un 59,70% que registró el modelo de Fulmer (59,70%).

Por lo que se refiere a la capacidad de predecir la bancarrota, el modelo (C7) se mostró superior al modelo (C13) para predecir la quiebra dos años antes (66,67% frente a 65,70%), pero su capacidad de predicción decae cuando lo hace a tres años (62,5% frente a 65%). Ambos modelos se mostraron claramente superiores a Fulmer et al. (1984).

Posteriormente, Wasilewski y Madra (2008) publicaron un trabajo que trataba de analizar el impacto que, en el índice Z de Altman, tenía la selección de una muestra con un conjunto de características representativas de la empresa agraria (propiedad legal de la tierra, área de cultivo, valor de la tierra, proporción del capital en el total de los pasivos, rentabilidad de la tierra, rentabilidad del capital) en Polonia. La muestra estaba compuesta por 123 empresas agrarias durante el periodo comprendido entre 2000 y 2004.

Las variables empleadas en la estimación del modelo fueron las del modelo de Altman. Se calcularon diferentes resultados para el índice Z, en base a una clasificación de la muestra de acuerdo con algunas de las citadas características específicas de la empresa agraria. De este modo, se definieron tres niveles para el índice Z: a) Aquellas empresas cuyo índice se situaba por debajo de 1,23, las cuales fueron clasificadas como insolventes; b) Aquellas empresas cuyo índice se situaba entre 1,23 y 2,90 y cuyo estado fue calificado como ambiguo; c) Aquellas empresas con un índice superior a 2,90, que fueron clasificadas como solventes.

De acuerdo con esta metodología, se obtuvieron una serie de valores para el Z-índice en función de los diferentes parámetros característicos de la empresa agraria. El Z-índice,

teniendo en cuenta la propiedad legal de la tierra, tomó valores en el rango de 1,23 y 2,90, situándose en una zona ambigua. El Z-índice, atendiendo al área de cultivo, arrojó valores medios, por lo general, de entre 1,76 y 2,32 ubicándose en una zona ambigua a pesar de que en algún año aparecieron valores extremos. Teniendo en consideración el valor de la tierra, el Z-índice tomó un valor predominante alrededor de 2,51, siendo ésta una zona ambigua también. El Z-índice, según el porcentaje de capital sobre el total de los pasivos, arrojó valores de 3,98 y 6,62 en algunos años. Cuando se tiene en cuenta la rentabilidad del capital, el Z-índice arrojó en promedio valores en la zona ambigua, si bien se obtienen mayores valores del índice para aquellas empresas con mayor porcentaje de capital. Si se tiene en consideración la liquidez como el ratio entre activos corrientes y pasivos corrientes, las empresas con menores ratios obtienen los menores índices. Por último, atendiendo a la rotación de proveedores, el Z-índice se situó en la zona de ambigüedad para la mayoría de los años.

D'Antoni et al. (2009) publicaron un estudio cuyo objetivo era investigar los factores que predecían el "estrés financiero", entendido como la insuficiencia de activos líquidos con los que atender las obligaciones financieras y no financieras de los granjeros jóvenes y noveles de Estados Unidos.

Para ello estimaron un modelo logit multinomial con una muestra compuesta por 19.638 granjas de una vida productiva inferior a los 10 años que se tomó del Agricultural Resource Management Survey (ARMS) para el periodo 2004-2006.

Dicha muestra se clasificó en cuatro grupos: 1) Granjas con una posición favorable (Deudas con un peso inferior al 40%

de sus activos e ingresos netos positivos); 2) Granjas con solvencia marginal (Ingresos netos positivos y elevada deuda); 3) Granjas con ingresos marginales (Resultados negativos y baja deuda) y 4) Granjas en posición financiera vulnerable (Deudas por encima del 40% de sus activos y resultados negativos).

Las conclusiones, que se centraron exclusivamente en los resultados del grupo de granjas en posición financiera vulnerable, mostraron que la edad del granjero, el tamaño de las operaciones, la propiedad, el año de operación y el tipo de granja son determinantes claves del estrés financiero, de acuerdo con la definición que del mismo ofrece el propio artículo.

Posteriormente, Susicky (2011) analizó los resultados de la aplicación de seis modelos de predicción de insolvencia en compañías agrícolas de la República Checa. Para la elaboración del estudio se acudió a diversas bases de datos: la base de datos Magnus Database Internet Securities y la Information Server of the Czech Judiciary.

Los diferentes modelos obtuvieron resultados dispares en relación con la predicción de la insolvencia en las compañías agrícolas de la muestra. Así, el modelo Zeta obtuvo un 94,6% de aciertos. El IN99 registró un 90,2% de exactitud en la predicción. El Z-score y el Z´score un 83,8% y un 60,4%, respectivamente. El IN05 y el IN01 obtuvieron un 74,0% y un 70,8%, individualmente. Finalmente, el modelo de Taffler registró una precisión del 54,1%.

Sin embargo, en conjunto, este trabajo ofrece una clasificación de los modelos atendiendo no sólo a aquellos que con más precisión predicen la insolvencia sino también a los que presentan menores errores tipo I y tipo II. Atendiendo a estos criterios de prelación, el orden de relevancia de los modelos es: Primero, el modelo Z-Score; segundo, el modelo Zeta; tercero, el modelo IN99; cuarto, el modelo Z´Score; quinto, el modelo IN05; sexto, el modelo Taffler y séptimo, el modelo IN01.

Mateos-Ronco et al. (2011) publicaron una revisión bibliográfica en la que se constató la falta de estudios sobre la predicción de la insolvencia en cooperativas agrarias. Allí se discuten aquellas aportaciones y estudios que, sin ser propios de cooperativas agrarias, pudieran ser de aplicación a esta tipología concreta de sociedades en España. A partir de esta revisión bibliográfica, seleccionan treinta y cinco variables financieras con potencial de aplicación a cooperativas agrarias.

Finalmente, en Vavrina et al. (2013) se comparan diversos métodos de predicción de la insolvencia en empresas agrarias. Se comparan tanto los métodos tradicionales y ampliamente conocidos (Z Score, MDA y Logit) como nuevos enfoques (DEA-Data Envelopment Analysis) y (PFEP-Production function-based economic performance evaluation). Los resultados muestran que todos los métodos tienen sus ventajas e inconvenientes. Sin embargo, se obtuvieron resultados prometedores cuando se aplicó DEA a una única empresa.

La base de datos que utilizaron fue Amadeus of Bureau van Dijk. De ella tomaron 2.581 empresas activas y 71 en quiebra de Checoslovaquia, Hungría, Polonia y Eslovenia, durante el periodo comprendido entre 1998 y 2002. Los

indicadores financieros empleados variaron según fuera el método empleado. Para Logit emplearon ocho variables. Para DEA se utilizaron nueve variables, y para el modelo de Altman se emplearon las cinco variables ampliamente conocidas.

Los resultados obtenidos fueron heterogéneos en función de la metodología empleada. Utilizando el Z-score de Altman, se observó que el porcentaje de acierto pasó del 62,1%, un año antes de la quiebra, al 57,8% cuatro años antes del evento. Mediante el empleo de Logit, el porcentaje de acierto pasó del 71,9%, un año antes de la quiebra, al 64,4% cuatro años antes. La utilización de DEA mostró que cuanto mayor es el número de empresas utilizadas en la muestra peor es la capacidad de esta metodología para predecir la insolvencia. Sin embargo, con una sola empresa quebrada en la muestra, DEA es capaz de detectarla perfectamente. Por último, mediante el empleo del método de la Frontera de Producción (PF), el porcentaje de aciertos resultó ser del 53,1% un año antes de la insolvencia, cayendo al 50,0% cuatro años antes de la quiebra. Sin tener en cuenta el periodo temporal, este método arrojó un acierto en la predicción del 62,9%.

1.4.2.2 Sector Industria

El objetivo del trabajo de Appetiti (1984) es desarrollar un modelo predictivo de la solvencia de las empresas italianas utilizando MDA, comparando la capacidad de este modelo, que usaba ratios estáticos, con otro donde estos ratios habían sido sustituidos por valores que tenían en cuenta la tendencia de los mismos. Los resultados obtenidos, considerando la tendencia de estos ratios durante tres años, fueron prometedores y mostraron cómo este modelo se comportaba mejor que el

primero, pero también mejor que cualquier otro modelo ya existente de empresas italianas.

Los datos usados en este trabajo provenían de un estudio llevado a cabo por el banco italiano Medio Banca durante los años 1979, 1980 y 1981. Incluía los datos financieros de 1.082 empresas italianas pertenecientes únicamente al sector manufacturero. El emparejamiento de empresas sanas y empresas en dificultades financieras se hizo atendiendo al subsector de actividad concreto, sin tener en cuenta el factor tamaño en dicho emparejamiento.

En el primer modelo, el ratio que mejor capacidad de discriminación mostró fue "Amortización acumulada/Activos Fijos", que clasificó correctamente 21/25 de las empresas en dificultades financieras (84%) y 20/25 de las empresas sanas (80%), mostrando una capacidad discriminatoria global del 82%. Por su parte, la función de dos variables que mejor resultado clasificatorio mostró añadió al ratio anterior el de "Fondo de maniobra/Deudas a largo plazo", que logró clasificar correctamente el 84% de las empresas. Además, se realizó un test de validación en una muestra separada y los resultados mostraron una tasa de error global del 30%.

En el segundo modelo se llevó a cabo una regresión para cada una de las 47 variables y para cada una de las 50 empresas. Entre las variables que mejor clasificación se obtuvo destaca "Gastos financieros/Valor añadido".

La principal aportación del trabajo de Appetiti fue la mejora del poder discriminatorio que proporcionaban los ratios financieros debido a la inclusión de variables de tendencia.

Zavgren (1985) desarrolló un modelo Logit de predicción de insolvencia para un periodo de cinco años previos al fracaso empresarial, junto con una metodología para la evaluación de la probabilidad de riesgo financiero. Para ello utilizó una muestra formada por 45 empresas sanas y 45 empresas quebradas durante el periodo 1972-1978, pertenecientes al sector manufacturero de Estados Unidos. Obtuvo una tasa de error global del 31%.

Theodossiou (1991) se propuso desarrollar un modelo fiable de predicción de la insolvencia para las empresas manufactureras griegas. Los datos financieros se recopilaban de los estados financieros de las empresas dos años antes de la quiebra. En concreto, la muestra de estimación se compuso finalmente por 54 empresas quebradas y 309 empresas sanas en el periodo 1980-1983, estando la muestra de testeo compuesta por 27 empresas quebradas y 111 empresas sanas en 1984.

Este estudio de Theodossiou (1991), con la muestra de entrenamiento, registró un porcentaje de error en la clasificación de empresas sanas/quebradas del 3,56% y del 7,41%, respectivamente. Con la muestra de testeo, el porcentaje de error en la clasificación de empresas sanas/quebradas fue del 8,11% y del 3,70% respectivamente.

El trabajo de Zhang et al. (1999) presentó un esquema general que permitía entender el papel de las NN en la predicción de la insolvencia. Realizó una revisión exhaustiva de las aplicaciones de las NN en este campo e ilustró su vínculo con la teoría de la clasificación Bayesiana. Dispuso de una muestra integrada por 110 empresas manufactureras

quebradas desde 1980 hasta 1991 y 110 empresas manufactureras sanas en el mismo periodo. Las empresas fueron ordenadas en función de los dígitos del SIC (Standard Industrial Classification) y del volumen de activos.

Los resultados fueron concluyentes en el sentido de que mostraron que NN son significativamente mejores que Logit, tanto en la predicción como en el porcentaje de clasificación. Para reforzar estos resultados recurrieron a la variación de las muestras, siendo en todas ellas las NN superiores al modelo Logit, con porcentajes que oscilan entre un 88,64% y un 86,36% en NN y entre el 80,68% y el 77,27% para Logit.

El trabajo de Sung et al. (1999) usó un enfoque de minería de datos para desarrollar modelos de predicción de insolvencia que fueran apropiados para circunstancias normales y para circunstancias de crisis económica. El modelo reveló que las principales variables que predicen la insolvencia eran "Cash flow/Total Activo", "Productividad del capital", en condiciones de normalidad, y "Activos Fijos/Fondos Propios + Pasivo a largo plazo", en condiciones de crisis. Las tasas de acierto del modelo en condiciones económicas normales y en condiciones de crisis fueron de un 83,0% y de un 81,0%, respectivamente.

Para el periodo de crisis, se tomaron todas las empresas que quebraron desde el segundo trimestre de 1997 hasta el primer trimestre de 1998, a partir de la información del KSE (Korea Stock Exchange). Inicialmente, 75 empresas fueron seleccionadas y, después de la aplicación de determinados filtros, quedaron 30 empresas del sector manufacturero. Para el periodo normal, inicialmente se seleccionaron 56 empresas cotizadas que después de ser filtradas quedaron en 29

empresas que entraron en quiebra. Como muestra de control, cada empresa en quiebra se emparejó con una o dos empresas sanas de acuerdo con el año, el tamaño y el número de empleados. Quedaron, finalmente, 49 y 54 empresas sanas para el periodo normal y de crisis.

El propósito del trabajo de Darayseh et al. (2003) era comprobar si un modelo Logit que utilizara variables macroeconómicas junto con ratios financieros produciría mejores resultados que los modelos que tan solo utilizaran ratios financieros. A partir de la base de datos COMPUSTAT se seleccionaron 110 empresas manufactureras quebradas de Estados Unidos durante el periodo 1990-1997, siendo emparejadas con 110 empresas manufactureras sanas. El emparejamiento se llevó a cabo teniendo en cuenta el código de actividad económica y el volumen de activos.

El modelo propuesto predijo correctamente el 87,82% de la muestra de entrenamiento y el 87,50% de la muestra de testeo, respectivamente. La significatividad de los coeficientes se midió empleando el estadístico "t" correspondiente a cada valor del coeficiente. Los modelos en su conjunto resultaron ser significativos al 5%.

Por su parte, el objetivo del trabajo de Yi-Chung y Fang-Mei (2005) fue comparar dos tipos de NN (MLP y RBFN) con las tradicionales técnicas de clasificación (MDA y Logit) en la predicción de insolvencia aplicada a empresas manufactureras del Reino Unido.

La muestra estaba compuesta por 904 empresas cotizadas pertenecientes al sector industrial general. De éstas,

353 quebraron o fueron adquiridas por otra empresa entre marzo de 1985 y marzo de 1994. Los datos de las empresas quebradas se tomaron de los informes financieros un año antes de la quiebra. Finalmente, se seleccionaron 32 empresas quebradas y 45 empresas sanas de las 551 inicialmente consideradas, dividiéndose la muestra entre una muestra de entrenamiento (80% de las empresas) y una muestra de testeo (20% de las empresas).

En comparación con MDA, el análisis Logit ofreció muy buenos resultados (78,36% en la muestra de entrenamiento y un 87,50% en la muestra de testeo), siendo un método al que los autores sugerían prestar mayor atención. Asimismo, parece que la RBFN es un método superior a los demás (81,96% en la muestra de entrenamiento y un 93,75% en la muestra de testeo). Concluyeron que MLP y RBFN son herramientas potentes debido a sus propiedades de aprendizaje no lineales y no paramétricas, y que muestran menos restricciones en el análisis de un menor número de casos.

En el trabajo de Gómez et al. (2008) se pretendía comprobar la estabilidad de varios modelos de predicción de insolvencia que incorporaban peculiaridades sectoriales, variables no financieras y una ampliación del periodo temporal. Se obtuvieron tres modelos aplicando Logit que pronosticaban con 1, 2 y 3 años de antelación la probabilidad de que la pequeña empresa industrial española se viera envuelta en un proceso concursal.

Tomando datos del Registro Mercantil, la muestra se compuso por 182 empresas del sector industrial (CNAE 151 a 410) que quebraron o suspendieron pagos en 2001, y 182

empresas industriales sanas. La información de las 364 empresas seleccionadas era de 1998, 1999 y 2000. A su vez, se obtuvo una muestra de validación para los mismos periodos, formada por 459 empresas, de las cuales 154 eran fallidas.

La clasificación de pertenencia de los tres modelos es muy satisfactoria, situándose los porcentajes obtenidos para las muestras de validación cercanos al 80% para 3 y 2 años antes de la quiebra, y al 83% para 1 año antes.

Habilollah y Farnoosh (2009) compararon la efectividad predictiva de NN y MDA. Los resultados obtenidos, a partir de los modelos basados en los datos de 80 compañías del sector manufacturero cotizadas en Teherán, mostraron que los modelos NN eran más precisos.

El objetivo del trabajo de Tzong-Huei (2009) fue comparar los resultados de cuatro técnicas de clasificación (MDA, Logit, Probit y NN) empleando datos de empresas industriales de Taiwán. Asimismo, se comparó la capacidad de los tres modelos más empleados (MDA, Logit y Probit) con datos de Taiwán y usando las variables de los trabajos clásicos de Altman (1968), Zmijewski (1984) y Ohlson (1980). La predicción se realizó para 1 y 3 años antes de la quiebra. Para este trabajo se utilizó una muestra de empresas industriales cotizadas de Taiwán durante el periodo 1998 a 2005, y quedó finalmente compuesta por 96 empresas quebradas y 158 empresas no quebradas. Los resultados, con la muestra de testeo, mediante MDA, mostraron que un año antes de la quiebra el modelo presentaba un error tipo I del 28,57% y un error tipo II del 16,67%. En el plazo de tres años, el error tipo I bajó al 12,5% y el error tipo II subió ligeramente al 30,0%.

El modelo Probit mostró que un año antes de la quiebra el modelo presentaba un error tipo I del 21,43% y un error tipo II del 8,33%. En el plazo de tres años, el error tipo I subió al 31,25% y el error tipo II escaló al 23,33%.

El modelo Logit reveló que, un año antes de la quiebra, el modelo presentaba un error tipo I del 21,43% y un error tipo II del 8,33%. En el plazo de tres años, el error tipo I subió al 37,5% y el error tipo II fue del 23,33%.

A su vez, el modelo NN mostró que un año antes de la quiebra el error tipo I fue del 6,84% y el error tipo II del 19,8%. En el plazo de tres años, el error tipo I subió al 21,43% y el error tipo II fue del 28,57%.

Los resultados del modelo, empleando las variables de Altman, evidenciaron que mediante MDA, un año antes de la quiebra y con la muestra de testeo, el modelo registró un error tipo I del 14,29% y un error tipo II del 24,0%. En el plazo de tres años, el error tipo I se disparó al 66,67% y el error tipo II al 46,88%. Con Logit, los porcentajes de error resultaron muy similares, siendo inferiores a los obtenidos mediante Probit.

Los resultados del modelo empleando las variables de Zmijewski mostraron que mediante MDA, un año antes de la quiebra y con la muestra de testeo, el modelo registró un error tipo I del 20,0% y un error tipo II del 29,63%. En el plazo de tres años, el error tipo I fue del 66,67% y el error tipo II del 46,88%. Con Logit, los porcentajes de error resultaron muy similares, siendo inferiores los obtenidos mediante Probit, al igual que en el caso del uso de las variables de Altman.

Los resultados del modelo empleando las variables de Ohlson mostraron que mediante el Análisis Discriminante, un año antes de la quiebra y con la muestra de testeo, el modelo registró un error tipo I del 20,0% y un error tipo II del 7,41%. En el plazo de tres años, el error tipo I se disparó al 60,0% y el error tipo II al 52,63%. Con Logit, los porcentajes de error resultaron muy similares, siendo inferiores a los obtenidos mediante Probit, de manera similar a lo ocurrido con el uso de las variables de Altman y Zmijewski.

El trabajo de Xu y Zhang (2009) investigó si se podía predecir la insolvencia en las empresas japonesas cotizadas desde 1992 a 2005, recurriendo a diversas metodologías (Z-score de Altman, O score de Ohlson, y la distancia a la quiebra basado en la teoría de opciones desarrollada en Estados Unidos). El poder predictivo de estas metodologías quedó reforzado mediante la combinación de dichos métodos. Adicionalmente, se construyó una nueva medida que incorporaba el grado de dependencia bancaria y del conglomerado conocido como Keiretsu.

Se utilizó la población de empresas japonesas cotizadas. La mayoría de las empresas cotizadas en Japón lo hacen en la Bolsa de Tokio, quedando la muestra compuesta por 3.510 empresas. La mayoría de las empresas pertenecía al sector manufacturero y se identificaron 76 empresas quebradas.

El modelo basado en la teoría de opciones es el que ofreció mejores resultados frente al resto basado en ratios contables. Cuando se combinaron ambos conjuntos de variables, los resultados mejoraron sustancialmente.

Al incorporar la variable “relación Banco/Keiretsu”, la predicción de la insolvencia mejoró.

Appiah y Abor (2009) usaron la información financiera de empresas manufactureras privadas medianas del Reino Unido durante el periodo 1994-2004, para determinar si el fracaso empresarial podía ser predicho mediante el desarrollo de un modelo Z-score. Después de aplicar diversos criterios, la muestra quedó compuesta por 31 empresas quebradas que fueron emparejadas con otras 31 empresas sanas. Con los mismos criterios, se extrajo una muestra de validación formada por 17 empresas quebradas y 20 empresas sanas.

La mayor contribución de este trabajo es haber puesto de manifiesto la superioridad, como variable, del “margen neto de explotación” en vez del “margen bruto de explotación” en el modelo empleado.

Con el objetivo de mejorar la precisión de las predicciones sobre insolvencia, Jae Kwon (2012) analizó los datos anuales de 1.888 empresas manufactureras recogidas por el Korea Credit Guarantee Fund (KODIT). Desarrolló un modelo de predicción basado en una función radial de máquinas de soporte (RSVM) y comparó la precisión de esta metodología con otras técnicas de inteligencia artificial.

Usó como muestra un conjunto de 944 empresas quebradas y 944 empresas sanas pertenecientes al sector manufacturero con un volumen de activos de entre 1 a 7 millones de dólares. El periodo considerado fue desde 1999 a 2005 y se crearon, para cada grupo, dos submuestras: una

muestra de entrenamiento (60%) y otra muestra de testeo (40%).

El trabajo comparó el rendimiento predictivo de la metodología RSVM frente a MDA, Logit, MLP, C5.0 y Redes Bayesianas. Midieron este rendimiento dividiendo el número de predicciones correctas entre el número total de predicciones. La comparación arrojó como resultados que RSVM obtuvo un valor medio de precisión del 82,35%; Logit mostró un valor medio del 78,45%; MLP dio por su parte un 78,31%; C5.0 mostró un 74,47%; y las redes bayesianas dieron el peor resultado, un 70,50%.

Los experimentos demostraron que RSVM siempre superó a otros modelos en la predicción de las dificultades financieras.

De Andrés et al. (2012) tratan el problema de la predicción de insolvencia con un nuevo procedimiento para mejorar el rendimiento de los predictores tradicionales (ratios). La propuesta consistía en sustituir dichos indicadores tradicionales por un análisis multinorma. Es decir, las desviaciones de cada empresa de un conjunto de indicadores sectoriales se usaban como nuevas variables predictoras. Los resultados indicaron que este enfoque podía mejorar significativamente la capacidad de predicción de los modelos.

La muestra utilizada para este estudio se dividió en dos grupos: En una primera muestra se tomaron las empresas que se declararon en quiebra desde julio de 2010 a junio de 2011 y con volumen de activos inferior a 10.000 euros. Al final, se seleccionaron 122 empresas (61 quebradas y 61 sanas) ordenadas por código de actividad y tamaño. Una segunda

muestra se compuso por 4.127 empresas necesarias para ajustar las curvas de cuantiles condicionados. Asimismo, la muestra fue dividida en una muestra de entrenamiento y en otra de testeo.

El análisis de ratios se basa en fuertes supuestos de linealidad y proporcionalidad. El enfoque de este trabajo, en el que se utilizan como variables independientes las desviaciones de los ratios observados con respecto a determinadas referencias de comportamiento en el sector (bajo rendimiento, rendimiento medio y alto rendimiento), mejora los resultados de la predicción en los tres modelos empleados (MDA, Logit y NN). Estos resultados podrían apuntar a que la mayoría de los trabajos que se han centrado en la mejora de las técnicas de predicción podrían estar produciendo resultados subóptimos.

Kwak et al. (2012) evaluaron la capacidad de las técnicas de minería de datos, mediante una técnica de programación lineal multicriterio, usando datos de empresas coreanas en quiebra justo después de la crisis de 1997. Los resultados mostraron que este método funcionaba tan bien como los métodos tradicionales al uso. Asimismo, los resultados eran comparables a los obtenidos con técnicas como SVM o DT, si bien no eran generalizables debido a que los datos procedían de empresas en un periodo temporal muy particular.

La muestra se compuso de 65 empresas quebradas y 130 empresas sanas del sector industrial durante los años 1997 a 2003. Los resultados con el modelo de Altman arrojaron un error global del 12,8%. Con el modelo de Ohlson, el error global fue del 16,4%. Y, finalmente, con una combinación del modelo de Altman y de Ohlson el error global fue del 11,3%.

El objetivo del trabajo de Callejón et al. (2013) era aumentar el poder predictivo de los modelos de NN para discriminar entre empresas solventes e insolventes incorporando para este propósito un nuevo conjunto de ratios financieros. El modelo propuesto consiguió predecir correctamente entre el 92,5% y el 92,1% de la muestra de entrenamiento y de la de testeo, respectivamente, usando información financiera de los dos años anteriores a la quiebra.

La muestra estaba compuesta por 1.000 empresas industriales europeas entre 2007 y 2009, emparejadas al 50%. Adicionalmente, se seleccionaron 140 empresas como muestra de testeo (70 solventes y 70 insolventes). Los emparejamientos se llevaron a cabo teniendo en cuenta el volumen de activos, el año considerado y el país.

Desarrollaron un modelo MLP basado en 6 variables independientes que consiguieron una precisión en la predicción del 92,11%, una sensibilidad del 94,69% y una especificidad del 89,66%. Los resultados revelaron que la solvencia depende básicamente de la capacidad de pagar la deuda con los fondos generados y del nivel de rentabilidad alcanzado por la propia compañía.

Zhang et al. (2013) usaron técnicas de fusión de información para construir un modelo de predicción anticipada de quiebra, basado en métodos de minería de datos, lo que posibilitaba integrar y aprovechar las fortalezas de los diferentes procedimientos para mejorar la capacidad de predicción. Se utilizó para ello una base de datos de empresas cotizadas manufactureras chinas a la que se aplicaron SVM y Logit. Los resultados mostraron una tasa de error global del

9,52% mediante SVM y del 9,52% para Logit. La combinación de ambas redujo la tasa de error al 4,76%.

Grüenberg y Lukason (2014) crearon modelos de predicción de insolvencia mediante NN y Logit, usando datos de empresas manufactureras estonias durante los años 2005-2008. Los resultados mostraron que métodos diferentes podían tener unas capacidades predictivas similares. Asimismo, se comprobó que la capacidad de predicción a 2 y 3 años vista resultaba peor. También indicaron que los modelos no funcionaban tan bien cuando se empleaban datos de empresas que se encontraban en plena crisis económica.

La muestra consistió en toda la población de empresas manufactureras estonias durante el periodo 2005-2008. Se utilizaron los datos de las quiebras ocurridas en 2009-2010 para chequear el poder predictivo de los modelos, tomándose un total de 16.360 empresas frente a 83 en quiebra (durante el periodo 2005-2008). Para la elaboración del modelo se emplearon 11.542 empresas frente a 58 en quiebra (2005-2008). Como muestra de testeo, se utilizaron 4.908 empresas frente a 25 en quiebra (2005-2008) y 8.990 empresas frente a 190 en quiebra (periodo de crisis 2009-2010).

La capacidad predictiva de los modelos mostró que, mediante Logit, el acierto era del 72% en el año previo a la quiebra, cayendo al 38% el tercer año. Mediante NN, los porcentajes de acierto fueron más estables, registrándose un 85% de acierto en el primer año y un 83% en el tercer año previo a la quiebra.

Bartoloni y Baussola (2014) proporcionaron un análisis del riesgo de crédito en las empresas manufactureras empleando los métodos MDA y DEA. La muestra principal se compuso por 5.000 empresas manufactureras italianas. De esta muestra, se extrajo una submuestra de 150 empresas quebradas durante los años 2005 y 2006. Los resultados de la combinación de ambos métodos (paramétrico y no paramétrico) arrojó las siguientes conclusiones: El 49% de las empresas en 2003 se encontraban correctamente clasificadas. Este porcentaje se elevó al 74,5% si se tenía en cuenta una clasificación menos restrictiva. De otro lado, el 51% de las empresas en 2009 se encontraban correctamente clasificadas de acuerdo a los mismos principios. Este porcentaje se elevó al 77,8% si se relajaba la clasificación. Finalmente detectaron que DEA funcionó mejor que MDA.

1.4.2.3 Sector Construcción

Langford et al. (1993) revisaron dos modelos para valorar la viabilidad financiera de las empresas constructoras. Discutieron la utilidad de las técnicas del análisis de ratios y del modelo Z para predecir si las empresas se dirigen a la insolvencia.

Llevaron a cabo el estudio de tres casos específicos de compañías pertenecientes al sector de la construcción, aplicando las técnicas mencionadas a empresas constructoras que hubieran quebrado en los últimos cuatro años. Emplearon seis ratios financieros: ratio de activos líquidos, test ácido, ratio Patrimonio Neto/Activos fijos, Capital circulante/Total activo, Beneficio después de impuestos/Patrimonio Neto, Beneficio después de impuestos/Valor empresa y Valor empresa/Total

pasivos. El modelo Z de Altman, por su parte, se aplicaba utilizando los cinco ratios habituales (Capital circulante/Total Activo, Beneficios retenidos/Total Activo, EBIT/Total Activo, Valor de Mercado de los Fondos Propios/Valor de Mercado de la Deuda a Largo Plazo y Ventas/Total Activo).

Concluyeron que, en tanto que las técnicas empleadas poseían valor a la hora de predecir algunas dificultades financieras, necesitaban ser usadas en conjunción con datos financieros de otras empresas de construcción. El artículo recomendaba la construcción de un modelo Z específico para las empresas constructoras.

Abidali y Harris (1995) trabajaron en el desarrollo de un sistema para la identificación de empresas constructoras en peligro de quiebra. El sistema se componía de un modelo predictivo discriminante que usaba ratios financieros, proporcionando un valor Z. Y de otro lado, de un sistema basado en la valoración de conductas directivas que generaba un valor específico A. El concepto que se hallaba detrás del valor A residía en la idea de que la compañía que se encuentra en dificultades financieras a menudo lo está debido a errores de actuación directiva previos.

Para la elaboración del Modelo Z se tomó una muestra de empresas constructoras británicas, con un número de empleados superior a 50 personas. Del total de 31 empresas seleccionadas, 11 de ellas habían entrado en quiebra entre 1978 y 1986. El resto, 20, eran empresas sanas en el periodo 1982-1986. Así mismo se utilizó un conjunto de 31 ratios financieros, consistentes en 24 ratios típicamente financieros y

7 medidas de tendencia. Finalmente, el modelo se estimó sobre 7 variables (4 ratios financieros y 3 variables de tendencia).

Por su parte, para la especificación del modelo A se describieron una serie de fallos comunes en las empresas que quiebran: a) La existencia de un CEO autocrático; b) Que la figura del Presidente y del CEO coincidan en la misma persona; c) La existencia de unos consejos de administración demasiado amplios; d) La ausencia de habilidades en ingeniería; e) Deficiencia de habilidades directivas; f) Unos sistemas de control insuficientes; g) Falta de experiencia en el mando y h) Pobres habilidades en marketing.

Otras variables que se consideraron importantes fueron el soportar demasiada deuda a corto plazo, presentar una tasa de crecimiento empresarial más acelerada que la tasa de crecimiento de la financiación necesaria, sufrir pérdidas en proyectos específicos y cometer errores importantes en la adquisición de empresas.

El modelo Z ofreció los siguientes resultados: a) El 90% de las empresas fueron correctamente clasificadas en el grupo de las no-quebradas, con un 10% clasificadas como vulnerables; b) El 100% de las 11 quebradas se clasificaron correctamente; c) El modelo también clasificó el 75% de las no quebradas como solventes en el año cuarto; d) El modelo clasificó el 73% del grupo en quiebra tres años antes del evento.

Validando el modelo en una muestra distinta, el 100% se clasificó correctamente como quebradas, el 67% se clasificó correctamente como sanas y el 20% como vulnerables.

Por lo que se refiere a los resultados obtenidos del modelo A, a partir del modelo Z y de sus resultados se construyeron dos grupos de 7 empresas cada uno, siendo un primer grupo solvente y un segundo grupo de empresas en riesgo. Se obtuvo un punto de corte mediante la suma de las ponderaciones asignadas a cada uno de los aspectos directivos referenciados en una encuesta. A partir de aquí se observó claramente cómo en el grupo de las de “en riesgo” registraban valores de A por debajo de 50. Asimismo, se observaron valores de A por encima de 50 para el grupo de las empresas “solventes”.

Mínguez (2006) aportó evidencia empírica sobre la previsión de la insolvencia en la construcción, dada la práctica inexistencia de modelos específicos para este sector. Llevó a cabo una estimación a través de un modelo Logit y de un modelo de riesgos proporcionales de Cox.

A partir de la base de datos SABI, y para el periodo comprendido entre noviembre de 1999 y diciembre de 2004, seleccionó 126 empresas con el código 45 del CNAE. De estas 126 empresas, 63 fracasaron y 63 se encontraban sanas. Finalmente, el número de casos válidos para el análisis se redujo a 26 empresas, siendo 13 las sanas y 13 las fracasadas.

Como predictores de insolvencia se eligieron inicialmente 32 ratios que, mediante la aplicación de técnicas de análisis univariante, quedaron reducidos a 16.

Los resultados del modelo Logit empleado se pueden resumir de la siguiente forma: a) 3,85% error tipo I y 19,23% error tipo II, con datos referentes al año inmediato anterior al

fracaso; b) 7,69% error tipo I y 11,54% error tipo II, con datos referentes al segundo año anterior al fracaso; y c) 11,54% error tipo I y 7,69% error con datos de tres años.

Por lo que se refiere a los resultados del modelo de Cox:

a) Un año antes, el error tipo I era 3,85% y 11,54% el error tipo II; b) Dos años antes, el error tipo I era 7,69% y el error tipo II 15,38%; c) Tres años antes, el error tipo I era 3,85% y el error tipo II 15,38%.

Stroe y Barbuta-Misu (2010) predijeron el resultado financiero de la empresa constructora en el condado de Galati (Rumanía), tomando del Registro de Comercio una muestra de 11 empresas constructoras para el periodo 2001-2006. Para llevar a cabo la separación entre empresas sanas y quebradas, estudiaron un conjunto de ratios financieros, de forma que se obtuviera un punto de corte. A partir de este análisis, se obtuvo: a) 7 empresas sanas a finales de 2006 y b) 4 empresas quebradas a finales de 2006. A partir de aquí, se calcularon unos coeficientes correctores y se calculó la función Z, estableciéndose una serie de rangos para clasificar el riesgo de quiebra. El modelo fue capaz de predecir la quiebra con un 81,82% de éxito.

Treewichayapong et al. (2011) analizaron la probabilidad de quiebra en una muestra amplia de empresas inmobiliarias en Tailandia, tanto cotizadas como no cotizadas, durante el periodo 2001-2009 mediante un modelo Logit y un modelo Cox de riesgos proporcionales. Comprobaron que los ratios financieros, como indicadores adelantados de quiebra empresarial, juegan un papel más destacado que las variables específicas de gobierno corporativo.

La muestra del estudio estaba compuesta por 96 empresas inmobiliarias. De estas, 32 estaban en quiebra y 64 estaban sanas. De las 32 en quiebra, 3 eran cotizadas y 29 no cotizadas. Y de las 64 sanas, 24 eran cotizadas y 40 no cotizadas. La muestra de empresas provenía del Ministerio de Comercio de Tailandia para las empresas no cotizadas, y de la Bolsa Tailandesa para las empresas cotizadas.

Estimaron un modelo Logit para el periodo 2001-2009 en el que los resultados de la estimación mostraron que la precisión del modelo a la hora de predecir la muestra de las empresas en quiebra, de las empresas sanas y de la muestra completa era del 94%, 97% y 96%, respectivamente. Por su parte, el error tipo I, es decir, aquel que clasifica erróneamente empresas en quiebra como sanas, fue del 6%.

Por lo que refieren como análisis de supervivencia, mostraron que la precisión del modelo para las empresas quebradas, no quebradas y la muestra completa fue del 66%, 94% y 84,4%, respectivamente, siendo estos porcentajes inferiores a los obtenidos mediante Logit.

En definitiva, la implicación de ambos modelos es que la empresa inmobiliaria Tailandesa que tiene un director con alta participación en el capital, un alto valor en el ratio "pasivos corrientes/total activos" y unas ventas bajas en relación con el inventario, presenta una alta probabilidad de quebrar.

Spicka (2013) comparó el resultado financiero de 81 empresas constructoras que quebraron durante el periodo comprendido entre enero de 2011 y mayo de 2013, con una muestra similar de empresas constructoras que gozaban de un

buen rating de crédito. El año empleado para la comparación fue el 2010.

La conclusión del estudio fue que una gestión financiera inadecuada es una de las fuentes principales del fracaso de la empresa constructora. Las compañías antes de la quiebra no gestionaron adecuadamente sus deudas en relación con su rentabilidad, en el sentido de que no generaron suficientes reservas para superar una crisis.

Para llegar a estos resultados emplearon el test no paramétrico de Mann-Whitney, que permite comparar las distribuciones de probabilidad de ambos grupos de empresas para diferentes predictores de fracaso empresarial.

Así, por lo que se refiere a los ratios de rentabilidad y productividad, muestra como el deterioro de las condiciones económicas afecta a ambas muestras. Sin embargo, aquellas empresas con menores reservas acumuladas se revelaron como las más vulnerables. En cuanto a la productividad, las empresas antes de la quiebra vieron reducidos drásticamente sus indicadores de productividad. Por lo tanto, el test de Mann-Whitney confirmó estadísticamente las diferencias en los indicadores de productividad y rentabilidad.

Con respecto a los predictores de insolvencia relacionados con la estructura de capital, la comparación entre ambas muestras reveló cómo la selección de empresas quebradas tenía una proporción de deuda sobre activos por encima del 100%. Se observó, por tanto, que el fracaso en las empresas de construcción estaba relacionado con un alto nivel de pasivos a corto plazo.

En relación con los ratios de liquidez y negocio, cabe señalar que estos resultaron significativamente menores en las empresas antes de la quiebra, momento en el que las empresas constructoras mostraron manifiestas dificultades de caja. Sin embargo, la rotación de activos no parecía diferir significativamente entre ambos grupos de empresas. En cambio, la rotación de activos y pasivos corrientes resultó ser significativamente inferior previamente a producirse el evento de la quiebra.

Finalmente, y por lo que se refiere a los ratios no financieros, la proporción de activos a largo plazo sobre los activos totales difirió significativamente entre ambos grupos. Las empresas constructoras antes de la quiebra mostraron una mayor proporción de activos a corto plazo.

En resumen, los resultados de los modelos de predicción sobre ambas muestras revelaron una precisión del 80% y del 82% para el grupo de empresas quebradas.

Gil de Albornoz y Giner (2013) investigaron si la estimación de modelos específicos sectoriales permitía predecir mejor el fracaso empresarial que la estimación de un modelo genérico.

Para la elaboración del estudio utilizaron una muestra compuesta por 4.600 empresas españolas, procedentes de la base de datos SABI, que presentaron concurso de acreedores durante el periodo 2005-2010. De esta muestra, la mitad aproximadamente pertenecían al sector constructor e inmobiliario, siendo el resto empresas industriales.

Para seleccionar los ratios que usaron como predictores de la insolvencia empresarial recurrieron a un procedimiento en dos etapas. Las autoras partieron de un conjunto de quince variables y, tras aplicarles diversos test, quedaron reducidas a cinco: a) Rentabilidad económica (Resultado de Explotación/Activo Total); b) Endeudamiento (Deuda total/Activo Total); c) Cobertura de intereses (Resultado neto sobre gastos financieros); d) Existencias sobre cifra de negocios; e) Liquidez (Activo corriente/Pasivo corriente).

Una vez seleccionados los predictores del riesgo de insolvencia, se estimó un modelo Logit para el periodo 2005-2009, empleando diferentes muestras pertenecientes a tres sectores económicos distintos (construcción, inmobiliario e industrial) y con una muestra general global.

Como conclusiones del trabajo destacaron que, en el sector de la construcción, la estimación sectorial permite clasificar mejor las empresas fracasadas que la estimación general (menor error tipo I, 22% frente a 32%), pero clasifica peor las empresas no fracasadas (mayor error tipo II, 45% frente a 33%). En conjunto, el error total de clasificación es superior en la estimación específica del sector (45%) que en la general (33%). En cambio, en el sector inmobiliario la estimación específica parece la más adecuada. Así, en la estimación específica se comete un error tipo II doce puntos porcentuales inferior al obtenido en la estimación general (41% frente a 53%).

En resumen, los resultados del trabajo confirmaron que la rentabilidad, la rotación de existencias, el endeudamiento, la liquidez, el crecimiento, la edad y las condiciones económicas

generales afectaban a la probabilidad de insolvencia, si bien pusieron de manifiesto la existencia de diferencias sectoriales.

1.4.2.4 Sector Comercio y Servicios

He y Kamath (2006) investigaron si dos modelos genéricos de predicción de insolvencia (Ohlson, 1980 y Shumway, 2001) mantenían su validez cuando eran aplicados a un sector en concreto, por ejemplo, el sector comercial. Los resultados sugirieron que la precisión de los modelos genéricos para clasificar se reducía dramáticamente cuando eran aplicados a muestras de un sector en concreto.

La muestra de este estudio fue de 354 empresas de diversos sectores, con un tamaño medio de activos inferior a los 37 millones de dólares. Incluyeron, asimismo, una submuestra de 40 empresas comerciales con activos inferiores a los 130 millones de dólares un año antes de la quiebra. El periodo de estudio se extendió desde 1990 hasta 1999. Finalmente, cada una de las 177 empresas quebradas se emparejó con una empresa sana del mismo sector, de parecido tamaño y del mismo año. Dentro de las variables empleadas, el modelo de Ohlson empleó 8 variables y el modelo de Shumway empleó 4 predictores para la insolvencia.

Los resultados obtenidos revelaron que, mediante el modelo de Ohlson, la capacidad para clasificar la muestra general mixta de sectores era del 88%, con un error tipo I del 12% y un error tipo II del 7%, un año antes de la quiebra. En cambio, la capacidad del modelo para clasificar correctamente a las 40 empresas comerciales era notablemente inferior, siendo

en concreto del 80%, 65% y 70%, un año, dos y tres años previos a la quiebra.

Mediante el modelo de Shumway la precisión en la muestra general era del 92%, 78% y 72%, para uno, dos y tres años previos a la quiebra. Para las empresas comerciales, la precisión sobre la muestra de las 40 empresas un año antes de la quiebra fue del 88%, si bien descendía por debajo del 60% dos años y tres años antes del evento de quiebra.

Keener (2013) presentó un modelo Logit de predicción de la insolvencia para una muestra de empresas del sector comercial durante el periodo 2005-2012. Los resultados muestran que las empresas comerciales pequeñas y con menos empleados tienen una mayor probabilidad de quiebra. Los resultados también aportan una fuerte evidencia de que las empresas con un menor ratio de caja sobre pasivos corrientes, menores márgenes de cash flow y mayores ratios de deuda sobre capital tienen mayor probabilidad de quebrar.

A partir de la base de datos COMPUSTAT, seleccionaron una muestra de 1.203 empresas en el periodo 2005-2012. Las variables empleadas en este estudio se seleccionaron atendiendo al criterio de que hubieran demostrado ser buenos predictores de la insolvencia en otros trabajos previos. En concreto, se seleccionó: 1) Número de empleados; 2) Rentabilidad sobre activos (ROA); 3) Margen de cash-flow; 4) Ratio deuda/capital; y 5) Ratio caja/pasivos corrientes.

Todas las variables del modelo resultaron ser significativas al 1%, excepto por lo que se refiere al ratio deuda/capital, que lo era pero al 10%. Los signos de las variables mostraron el

sentido esperado en todos los casos, salvo para el ratio ROA. El modelo presentó un R^2 del 91,7%. En resumen, los resultados demostraron que cuanto menor es el tamaño de las empresas y menor es el número de empleados mayor es la probabilidad de quiebra. Sin embargo, el modelo arrojó un resultado inesperado por cuanto relacionó una mayor probabilidad de quiebra con un mayor ROA.

1.4.2.5 Sector Hostelería

Dentro de la literatura sobre predicción de insolvencia son muy escasos los estudios realizados específicamente para el sector de la hostelería. El sector de hostelería comprendería actividades diversas que abarcarían desde los servicios de alojamiento pasando por los servicios de comida y bebida. Por tanto, se trataría de dos industrias diferenciadas aunque muy estrechamente relacionadas, como la de hoteles y de restaurantes, conocidas en la literatura anglosajona bajo el término "Hospitality" que, además, incluiría empresas dedicadas al servicio de diversión y entretenimiento.

Olsen et al. (1983) fueron los primeros en predecir la quiebra empresarial en la industria de restaurantes. Emplearon una muestra formada por 7 restaurantes en quiebra y 12 solventes. Utilizaron un análisis gráfico de la aplicación de distintos ratios contables en contraposición a los modelos estadísticos más sofisticados, lo cual fue considerado como la principal limitación del estudio. No obstante, su fácil aplicación en la vida real fue la principal ventaja aportada por esta investigación.

Cho (1994) estimó modelos Logit para predecir la quiebra de restaurantes y hoteles. Según el estudio realizado, fueron dos las variables que finalmente resultaron explicativas, en el caso de los restaurantes, alcanzando un nivel de predicción del 91% un año antes de la quiebra. Por otro lado, en el caso de los hoteles, tan solo resultó una variable explicativa, clasificando el modelo correctamente en el 92% de las empresas consideradas. Cabe destacar que en este estudio se consideraba que una empresa entraba en quiebra si había obtenido pérdidas en los tres últimos ejercicios.

Gu y Gao (2000) estimaron un modelo MDA de predicción de insolvencia basado en una muestra de 14 empresas en quiebra, incluyendo 4 hoteles y 10 restaurantes, y un número similar de empresas no quebradas. Los datos de las empresas quebradas correspondían al periodo 1987-1996. Seleccionaron cinco variables (Total Pasivo/Total Activos; Beneficio antes de intereses e impuestos [EBIT]/Pasivo corriente; Margen de beneficio bruto; Deuda a largo plazo/Total Activo; e Ingresos/Activos no corrientes). El modelo pudo clasificar las empresas, dentro de la muestra, con una precisión del 93%.

Gu (2002) utilizó una muestra de tamaño algo superior, considerando 18 restaurantes en quiebra, emparejados con otros tantos restaurantes no quebrados, con datos del periodo 1986-1998. Las empresas fueron emparejadas por año y por total activo. Gu desarrolló un modelo MDA exclusivamente para la predicción de restaurantes, siendo la muestra un conjunto de empresas americanas que cotizaban o habían cotizado en bolsa. Inicialmente consideró un total de 12 ratios, resultando finalmente explicativos en su modelo, con un nivel de acierto del 92% un año antes de la quiebra, tan solo dos de ellos

(EBIT/Total Pasivo y Total Pasivo/Total Activo). El modelo estimado se aplicó posteriormente a una muestra externa, formada por 49 empresas no incluidas en la muestra inicial (48 empresas solventes y 1 empresa insolvente), obteniéndose un nivel de predicción fuera de la muestra del 80% (ligeramente superior a la capacidad de predicción del modelo de Altman (1968) fuera de la muestra, que se situó en un 79%).

Los primeros en utilizar un modelo Logit aplicado a este sector fueron Kim y Gu (2006a), con una muestra de 32 empresas del sector hostelería (16 en quiebra y otras 16 no quebradas). Estimaron dos modelos Logit para predecir la quiebra 1 y 2 años antes de la misma, de empresas pertenecientes al sector hotelero, restaurantes y empresas que prestan servicios de entretenimiento. Este estudio tuvo como aportación principal la estimación de un modelo para predecir la quiebra 2 años antes, a diferencia de los anteriores estudios predictivos del sector que tan solo lo habían hecho 1 año antes de la quiebra. De los 13 ratios inicialmente considerados, resultó significativamente explicativo uno de ellos: Cash-flow operativo/Total Pasivo. La capacidad predictiva del modelo se situó en el 91% para 1 año antes de la quiebra y en el 84% para 2 años antes. Los autores concluyeron, en base a los modelos estimados, que las empresas de este sector están más próximas a la quiebra si tienen un menor flujo de efectivo y un mayor endeudamiento.

Posteriormente Kim y Gu (2006b) estimaron un modelo Logit utilizando los mismos datos del modelo MDA de Gu (2002). Al utilizar los mismos datos se pudo comparar la capacidad predictiva de ambos modelos. El modelo Logit, que suponía una mayor sofisticación estadística, predecía

correctamente un año antes de la quiebra el 94% de las empresas que habían entrado en bancarrota, así como el 93% de las empresas utilizadas en la muestra externa. Con estos resultados se demostraba que el modelo Logit aportaba una capacidad predictiva algo superior (94% frente al 92% del MDA).

En ambos estudios se partía de un total de 12 ratios financieros que abarcaban tanto variables de solvencia como de rentabilidad y de eficiencia. Finalmente ambos modelos consideraron sólo dos de estas variables como las que mejor predecían la quiebra de los restaurantes: Total Pasivo/Total Activo y EBIT/Total Pasivo. En base al modelo estimado, Kim y Gu (2006b) concluyeron que las empresas de restaurantes con bajos beneficios antes de intereses e impuestos (EBIT) y alto endeudamiento son más proclives a entrar en quiebra.

Youn y Gu (2010a) realizaron un estudio empírico posterior que se centró en la industria de restaurantes utilizando información de empresas que cotizaban o habían cotizado en bolsa. Estos autores fueron los pioneros en aplicar NN en la predicción para esta industria. Conjuntamente estimaron un modelo Logit con objeto de comparar la capacidad predictiva de ambas técnicas, estimando sendos modelos para predecir la quiebra 1 y 2 años antes de la misma.

Consideraron un total de 10 variables iniciales, resultando finalmente dos con suficiente relevancia explicativa. Este estudio se centró en restaurantes americanos, con una muestra de 31 restaurantes quebrados durante el periodo 1996-2008, de los cuales 21 formaron parte de la muestra utilizada para predecir el modelo y los 10 restantes fueron utilizados para la

muestra externa que serviría como comprobación del poder predictivo. Igualmente se utilizaron otras 31 empresas de restaurantes no quebradas, realizándose un emparejamiento de la muestra por tamaño (utilizando como medida del tamaño el total activo). Al comparar la capacidad predictiva de los dos modelos, determinaron que, si bien NN predicen bien, no aportan mejor predicción que el análisis Logit, especialmente en la muestra externa. Al respecto, Charitou et al. (2004) mostraron que ambos podían utilizarse indistintamente y no daban un mejor resultado. Además, Logit permite interpretar los coeficientes de las variables de los modelos y cómo afectan a la probabilidad de quiebra, señalando vías para evitarla.

A este trabajo le siguió otro, realizado por los mismos autores, Youn y Gu (2010b), que se centró exclusivamente en empresas hoteleras coreanas, estimando igualmente un modelo Logit y otro NN. Obtuvieron una mejor capacidad predictiva al aplicar NN, tanto dentro de la muestra como en el testeo realizado con una muestra externa (83,33% frente al 77,27% obtenido en el modelo Logit dentro de la muestra y el 87,75% frente al 81,82% obtenido en la muestra de testeo). Si bien los resultados eran superiores, reconocían la limitación de los modelos NN en tanto no identifican las variables determinantes de la quiebra, y en consecuencia, no facilitarían información de los factores a mejorar para intentar evitarla. Asimismo constataron la importancia de la cobertura de intereses para mantener la solvencia de las empresas de alojamiento coreanas.

Youn y Zheng (2010) elaboraron un estudio en el que, usando datos financieros de empresas de restauración en Estados Unidos, desarrollaron dos modelos de predicción de

insolvencia con Logit y NN. Sus resultados revelaron que Logit no resultaba ser inferior a NN en términos de precisión en la predicción. En concreto, y para el caso de las empresas de restauración, el modelo logístico resultó útil a la hora de indicar qué deben hacer las empresas para reducir la probabilidad de quebrar.

A partir de la muestra inicial de 31 empresas de restauración, se tomaron 10 empresas quebradas con sus correspondientes 10 empresas sanas como muestra de validación. Para la muestra de entrenamiento se emparejaron un total de 21 empresas de restauración quebradas con otras 21 empresas sanas. Por último, se emplearon 10 ratios financieros como variables-predictores del fracaso empresarial.

Los resultados del estudio mostraron que, mediante el uso de Logit, el modelo produce una precisión en la predicción global del 88,1%, un año antes del evento de la quiebra, y de un 76,19% dos años antes. La capacidad predictiva del modelo en la muestra de validación arrojó un 95% de precisión un año antes y un 85% de precisión dos años antes de la quiebra. Ambos modelos no presentaron errores tipo I. Por el contrario, presentaron un error tipo II del 10% en el modelo de predicción un año antes de la quiebra y del 30% en el modelo de predicción dos años antes de la quiebra.

Por lo que se refiere a los resultados obtenidos mediante modelos NN, un año antes de la quiebra, el modelo predijo correctamente un 88,1%, en tanto que dos años antes, la capacidad predictiva descendió al 78,57%. La capacidad de predicción sobre la muestra de validación fue del 95%, un año antes de la quiebra, y del 80%, dos años antes de dicho evento.

Kim (2011) llevó a cabo un estudio con dos objetivos. Por una parte, proporcionar una aproximación óptima a la predicción de insolvencia de los hoteles que minimizara el riesgo de una mala clasificación y, por otra, investigar las características funcionales de MDA, Logit, NN y SVM en la predicción de la insolvencia en hoteles. El trabajo concluyó que la mejor técnica de alarma preventiva a la hora de predecir la quiebra era SVM.

Para elaborar el trabajo, utilizaron una muestra construida con los datos financieros de 33 hoteles turísticos coreanos durante el periodo 1995-2002. La muestra estaba compuesta por los datos financieros de tres años previos a la quiebra de 18 hoteles, los correspondientes datos financieros de dos años previos a la quiebra de 12 hoteles y los datos financieros de un año antes de la quiebra de 3 hoteles. Se emplearon en total 17 predictores de insolvencia.

Los resultados del trabajo mostraron la precisión de los distintos métodos empleados. Así, mediante el empleo de MDA obtuvieron un error tipo I y un error tipo II del 47,6% y del 7,3%, respectivamente, registrándose un error global del 27,4%. La utilización de Logit mostró, en general, errores más elevados o, una menor precisión a la hora de la predicción de la insolvencia. Arrojó unos errores tipo I y tipo II del 22% y del 18,4%, respectivamente, y un error global del 20%. Por lo que se refiere a la precisión que ofrecían NN, estas generaron unos errores tipo I y tipo II del 4,8% y del 12,1%, en tanto que el error global fue del 8,4%. Finalmente, SVM fue la que registró el menor error global de predicción con un 4%, y el segundo mejor resultado, dentro de los métodos empleados, por lo que

se refiere a los errores tipo I y II, con un 6,8% y un 1,3% respectivamente.

Más actualmente, Park y Hancer (2012) estimaron modelos Logit y NN como métodos para predecir la insolvencia de hoteles y restaurantes. Utilizaron datos correspondientes al periodo 1990-2009 considerando una muestra de 80 empresas (40 quebradas y 40 no quebradas). Las empresas elegidas correspondían a empresas que cotizaban o habían cotizado en bolsa en ese periodo. NN obtuvo un mayor nivel de precisión en la predicción que Logit dentro de la muestra (97.5% frente al 90%). Sin embargo, cuando se realiza el test para su verificación en la muestra externa (formada por 8 empresas solventes y 8 insolventes que no se habían incluido en el modelo inicial), ambos modelos predicen en un 100%. Para el estudio consideraron inicialmente un total de 16 ratios financieros, resultando finalmente explicativos, en los modelos estimados, 3 de ellos (Fondo de Maniobra/Total Activo, Total Pasivo/Patrimonio Neto y Total Pasivo/Total Activo).

Este estudio resalta que, de los ratios considerados, el ratio "Total Pasivo/Total Activo" era una variable significativa, basado en los resultados del test "t" y del análisis Logit. Los autores consideraron además que, en tanto que es conocido que las empresas de este sector se encuentran altamente endeudadas, la conclusión que se podría extraer es que una holgada financiación externa, cuando no está acompañada por una adecuada rentabilidad de las inversiones, podría ser un motivo importante para llevar a las empresas a entrar en quiebra.

1.4.3 ALGUNAS CONCLUSIONES ACERCA DEL DESARROLLO DE LA LITERATURA SOBRE PREDICCIÓN DE LA INSOLVENCIA

La literatura sobre predicción de la insolvencia viene desarrollándose con fuerza desde la década de los sesenta del siglo pasado. Desde entonces, han venido apareciendo estudios acerca de la predicción del fracaso empresarial ante la necesidad, cada vez más apremiante, de disponer de herramientas útiles con las que poder diagnosticar la salud financiera de la empresa y su eventual deriva hacia la quiebra.

La perspectiva de nuestra revisión se ha orientado en analizar la literatura atendiendo a si los trabajos habían producido modelos descentrados o centrados. Es decir, si los modelos habían sido elaborados con muestras multisectoriales o bien, si estos modelos habían sido diseñados con muestras de un único sector o subsector de actividad.

A partir de la revisión efectuada, se observan, en general, varias tipologías de trabajos:

De una parte, hay estudios que tienen por finalidad construir un modelo de predicción para pronosticar el fracaso empresarial y que permiten clasificar a las empresas como quebradas o sanas. Estos son, con diferencia, la gran mayoría de trabajos (un resumen de los trabajos empíricos focalizados en modelos descentrados aparece en la tabla 1.1). Dentro de este grupo, podríamos añadir aquéllos cuyo modelo se propone servir de “señal de alarma” anticipando el evento de la quiebra.

Existe también un importante número de trabajos que se han centrado en replicar, con variaciones, los modelos más clásicos. Dentro de esta tipología podríamos destacar todos aquellos trabajos que replicaron los estudios de Altman, Ohlson, Zmijewski y otros, introduciendo diferentes modificaciones en las metodologías, muestras, periodos temporales y tratamiento de la información.

Asimismo, tampoco resultan despreciables en número los trabajos que se ocuparon de comparar distintos métodos de predicción para ver cuál de ellos pronosticaba con mayor precisión la insolvencia. Particularmente abundantes en los últimos años, son los estudios que trataron de mejorar los porcentajes de acierto generales con la introducción de técnicas cada vez más sofisticadas, llegando a emplear algoritmos matemáticos con los que combinar dos o más metodologías en un intento de arañar algunos puntos porcentuales en las tasas de acierto. A este tipo de trabajos los podríamos denominar de "búsqueda de precisión".

En los últimos años han venido apareciendo también modelos centrados en la predicción de insolvencia en un sector o subsector de actividad específico, en la búsqueda de factores concretos y determinantes a la hora de aproximar la predicción del fracaso en empresas de un sector en particular. Estos modelos han sido objeto de análisis en el apartado correspondiente a los modelos centrados (un resumen de los trabajos empíricos focalizados en modelos centrados aparecen en las tablas 1.2, 1.3, 1.4, 1.5 y 1.6).

Y, finalmente, aquellos trabajos que trataron de validar hipótesis específicas como pueden ser los trabajos de Libby

(1975), Norton et al. (1979), Aharony et al. (1980), Zimmer (1980), Zmijewski (1984), Levitan y Knoblett (1985), Casey y Bartczak (1985), Gombola et al. (1987), Hopwood et al. (1989), Aziz y Lawson (1989), entre otros.

Estas cinco tipologías de estudios podrían quedar reducidos a dos perspectivas básicas a la hora de abordar un estudio sobre predicción de la insolvencia, si atendemos a lo sugerido por Darayseh et al. (2003) quienes afirmaron:

“Los investigadores tienen diferentes aproximaciones a la hora de estudiar la quiebra o el fallo empresarial. Básicamente, existen dos aproximaciones: por un lado, el objetivo del modelo (o del constructor del modelo) puede ser predecir qué empresas terminarán quebrando en un periodo determinado de tiempo. En este contexto, Dietrich (1984) propuso que, entonces, el objetivo del investigador es desarrollar un modelo con una alta capacidad de predicción, con independencia de las variables utilizadas o de los valores de las variables. En este sentido, si la capacidad predictiva es el objetivo del modelo en cuestión, el valor de los coeficientes, su nivel de significación e incluso hasta sus signos no tienen apenas interés. Por lo tanto, la construcción del modelo se convierte en un mero ejercicio de ajuste de los datos.

Por otro lado, el investigador puede estar interesado en comprobar unas hipótesis. En este contexto, la selección de las variables, el sesgo en la estimación, la selección de la muestra y la significación estadística de los coeficientes de las variables son comparativamente más importantes. Por ejemplo, el estudio de Ohlson (1980) pretendió comprobar si determinados ratios financieros y otros factores guardaban alguna relación

con la quiebra de las empresas. Ohlson no sólo estaba preocupado con la capacidad predictiva del modelo sino también con los valores de los coeficientes que el modelo presentaba”.

Desde este punto de vista que nos ofrecen Darayseh et al. (2003), es posible constatar que el avance de la literatura ha seguido un patrón desequilibrado en la medida en que el número de trabajos que se han desarrollado, introduciendo metodologías estadísticas y computacionales cada vez más complejas y sofisticadas, para la búsqueda de porcentajes de acierto cada vez más altos, ha superado con creces en número e importancia a aquellos trabajos que se han enfocado hacia el contraste de hipótesis específicas con las que esclarecer un aspecto particular del fracaso empresarial como asunto de investigación.

En consecuencia, se tiene la impresión personal que, a partir de los primeros trabajos originales, los estudios posteriores se desarrollaron buscando el cauce más inmediato y disponible, capitalizando sin duda el importante avance de los métodos estadísticos y computacionales, lo que explicaría el desequilibrio encontrado entre los trabajos de “búsqueda de precisión” y de comparación de modelos frente a aquellos que contrastaron hipótesis concretas. Sin duda, la dificultad de elaborar una teoría comprensiva del fracaso empresarial orientó los esfuerzos hacia esta vertiente más práctica e inmediata.

Sin embargo, como se ha reseñado anteriormente, los métodos más novedosos o innovadores no garantizan unos mejores resultados porque, si bien a medida que estos métodos se han ido sofisticando, permitiendo errores muy pequeños y

residuales, la horquilla de fluctuación de estos errores se ha ampliado cada vez más.

Por tanto, y visto desde una vertiente eminentemente práctica, una vez alcanzado un porcentaje de acierto significativo y elevado, ¿cuál es el sentido de insistir en la investigación de métodos que nos den apenas un porcentaje mínimo de mejora, sin otras ventajas añadidas como la estabilidad en el periodo de estimación y la aplicabilidad en otros periodos temporales, sectores o regiones, por citar solo algunas?.

No podemos obviar que el desarrollo de los modelos descentrados se debió a la dificultad de conformar muestras sectoriales representativas en los modelos primigenios. Es decir, estos modelos se hacían con la información disponible en cada momento. Por ello, la mejora en la información disponible y en las bases de datos existentes debería animar al desarrollo de modelos sectoriales específicos.

Diversos autores ya afirmaron que las particularidades sectoriales añaden mejoras significativas en la explicación que ofrecen los modelos, pero esta cuestión no ha sido abordada de manera directa con un diseño ad-hoc para tal objetivo.

Otro aspecto en el que se debería de profundizar es en el desarrollo de modelos de estrés financiero basado en varias etapas, de tal forma que fuera de utilidad para una detección temprana de problemas que hiciera posible implementar medidas correctoras.

Los modelos desarrollados predominantemente en la literatura han sido modelos de predicción del fallo empresarial, eso sí, con enorme acierto en el pronóstico un año antes del evento de quiebra y con un acierto decreciente a medida que se amplía el plazo temporal de predicción. Sin embargo, con tal plazo de efectividad es difícil que tales modelos sean de ayuda para corregir rumbos o modificar aspectos esenciales de la marcha de una compañía en lugar de certificar el avance hacia un desenlace fatal.

Desde este punto de vista y teniendo en cuenta el patrón seguido por la literatura, en nuestro modesto entender, la agenda de la investigación en este campo debería orientarse, si atendemos a lo sugerido por Darayseh et al. (2003), hacia el desarrollo de estudios que buscaran el contraste de hipótesis concretas y que nos acercaran a un entendimiento mejor y más global del fracaso empresarial como problema.

Este trabajo pretende aportar un modesto grano de arena en este campo de la formulación de hipótesis. Y para ello retoma una cuestión que fue aparcada hace un tiempo, de alguna manera, debido a razones de diferencias en las muestras, periodos temporales u otras, como es la de comprobar con una sola muestra y en un periodo temporal concreto si los modelos centrados predicen mejor que los descentrados o viceversa.

Tabla 1.1. Principales modelos descentrados de predicción de insolvencia

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Beaver (1966)	158 empresas	Industrial amplio	A. Univariante	EE.UU.	<u>Muestra</u> : 90-92%		
Altman (1968)	66 empresas	Industrial amplio	MDA	EE.UU.	<u>Muestra</u> : 94%	72%	48%
					<u>Testeo</u> : 79%	-	-
Deakin (1972)	158 empresas	Industrial amplio	Test de clasificación	EE.UU.	<u>Quebradas</u> : 77%	96%	94%
	64 empresas		Dicotómica		<u>Sanas</u> : 82%	92%	82%
			MDA				
Ohlson (1980)	2.163 empresas	Industrial amplio	LOGIT	EE.UU.	<u>Muestra</u> : 96,1%	95,5%	
Casey y Bartczak (1985)	310 empresas	Industrial amplio (Cotizadas)	MDA	EE.UU.	<u>Quebradas</u> : 57%-90%	<u>Sanas</u> : 47%-87%	
			LOGIT		13%-63%	95%-98%	
Gombola et al. (1987)	154 empresas	Manufacturero/ Comercial	MDA	EE.UU.	<u>Muestra</u> : 85-89%	67-70%	78%
Hing-Ling Lau (1987)	400 empresas	Industrial amplio	MDA	EE.UU.	<u>Muestra</u> : 98%	96%	95%
					<u>Testeo</u> : 83%	92%	91%

**Tabla 1.1. Principales modelos descentrados de predicción de insolvencia
(Continuación)**

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Aziz et al. (1988)	98 empresas	Industrial amplio	MDA LOGIT	EE.UU.	<u>Muestra</u> : 88,8% <u>Muestra</u> : 91,8%	80,6% 84,7%	72,5% 78,6%
Odom y Sharda (1990)	129 empresas	Industrial amplio	NN	EE.UU.	<u>Muestra</u> : 100% <u>Testeo</u> : 81,8%		
Koh y Killough (1990)	105 empresas	Industrial amplio (Cotizadas)	MDA	EE.UU.	<u>Quebradas</u> : 78,6% <u>Sanas</u> : 88,6%		
Gilbert et al. (1990)	380 empresas	Industrial amplio	LOGIT	EE.UU.	<u>Quebradas</u> : 62,5% <u>Sanas</u> : 97,9%		
Coats y Fant (1992)	94 empresas	Industrial amplio	NN	EE.UU.	<u>Muestra</u> : 95%	92,9%	86,2%
Fletcher y Goss (1993)	32 empresas	Industrial amplio	LOGIT NN	EE.UU.	<u>Muestra</u> : 71,3% <u>Muestra</u> : 80,5%	82,4%	75,0%
Johnsen y Melicher (1994)	660 empresas	Industrial amplio	Modelo Binomial	EE.UU.	<u>Quebradas</u> : 77,68% <u>Sanas</u> : 95,56%		
			Modelo Multinomial		<u>Quebradas</u> : 94,64% <u>Sanas</u> : 98,29%		

**Tabla 1.1. Principales modelos descentrados de predicción de insolvencia
(Continuación)**

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Altman et al. (1994)	808 empresas	Industrial amplio	MDA NN	ITALIA	<u>Testeo</u> : 95,1% <u>Testeo</u> : 95,3%		
Charalambous et al. (2000)	278 empresas	Industrial amplio	LOGIT NN (BP)	EE.UU.	<u>Testeo</u> : 77,9% <u>Testeo</u> : 73,3%	68,6% 68,6%	64,0% 67,4%
Shumway (2001)	300 empresas	Industrial amplio (Cotizadas)	Hazard Model	EE.UU.	<u>Testeo</u> : 75%		
Yi-Chung et al. (2005)	77 empresas	Industrial amplio	NN (MLP) NN (RBFN) MDA PROBIT	UK	<u>Testeo</u> : 83,75% <u>Testeo</u> : 93,75% <u>Testeo</u> : 80,0% <u>Testeo</u> : 87,50%		
Tzong-Huei (2009)	254 empresas	Industrial amplio (Cotizadas)	MDA LOGIT PROBIT NN	TAIWAN	<u>E.Tipo I</u> : 28,57% <u>E.Tipo I</u> : 21,43% <u>E.Tipo I</u> : 21,43% <u>E.Tipo I</u> : 6,84%		12,5% 37,5% 31,25% 21,43%
Shuk-Wern et al. (2011)	105 empresas	Industrial amplio (Cotizadas)	LOGIT	MALASIA	<u>Testeo</u> : 90%		

**Tabla 1.1. Principales modelos descentrados de predicción de insolvencia
(Continuación)**

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Chen et al. (2011)	10.366 empresas	Industrial amplio	SVM LOGIT		<u>Muestra:</u> 68,5% <u>Muestra:</u> 64,5%		
Kwak et al. (2012)	195 empresas	Industrial amplio	PROGRAMACIÓN LINEAL MULTICRITERIO	COREA DEL SUR	<u>Muestra:</u> 88,70%		
Sangjae y Wu Sing (2013)	229 empresas	Industrial amplio (Cotizadas)	NN (BP) MDA	COREA DEL SUR	<u>Muestra:</u> 81,43% <u>Muestra:</u> 74,82%		

NOTA: Elaboración propia.

Tabla 1.2. Principales modelos centrados de predicción de insolvencia (Sector Agricultura)

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Dietrich et al. (2005)	67 empresas	Agricultura	MDA	CANADÁ	M. C7: 66,6%-62,5%		
					M. C13: 65,7%-65,0%		
Susicky (2011)	496 empresas 974 empresas	Agricultura	ZETA, Z-SCORE, Z'-SCORE, TAFFLER y TISSHAW, IN99, IN01, IN05	REPÚBLICA CHECA	<u>Zeta</u> : 94,6%		
					<u>IN99</u> : 90,2%		
					<u>Z-score</u> : 83,8%		
					<u>IN05</u> : 74,0%		
					<u>IN01</u> : 70,8%		
					<u>Z'Score</u> : 60,4%		
					<u>M. Taffler</u> : 54,1%		

Tabla 1.2. Principales modelos centrados de predicción de insolvencia (Sector Agricultura) (Continuación)							
AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Vavrina et al. (2013)	2.652 empresas	Agricultura	Z-SCORE LOGIT DEA	CHEQUIA HUNGRÍA POLONIA ESLOVENIA	Z-Score (Altman)	62,1%	57,8%
					Regresión Logística	71,9%	64,4%
					Los resultados de este enfoque muestran que a mayor número de empresas quebradas en la muestra, PEOR predice DEA. Sin embargo, con una sola empresa quebrada en la muestra, DEA la detecta perfectamente.		
					Frontera de Producción	53,1%	50,0%

NOTA: Elaboración propia

Tabla 1.3. Principales modelos centrados de predicción de insolvencia (Sector Industria)

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Appetiti (1984)	1.082 empresas	Industrial Manufacturero	MDA	ITALIA	<u>Muestra:</u> 92,0% <u>Testeo:</u> 84,0%		
Zavgren (1985)	90 empresas	Industrial Manufacturero	LOGIT	EE.UU.	<u>Muestra:</u> 69,0%		
Theodossiou (1991)	363 empresas	Industrial Manufacturero	MLP LOGIT PROBIT	GRECIA	<u>Muestra:</u> 87,1% <u>Muestra:</u> 94,0%		
Zhang et al. (1999)	220 empresas	Industrial Manufacturero	NN LOGIT	EE.UU.	<u>Muestra:</u> 86,4%-89,2% <u>Testeo:</u> 85,0%-87,7% <u>Muestra:</u> 77,3%-80,7% <u>Testeo:</u> 76,8%-80,0%		
Darayseh et al. (2003)	220 empresas	Industrial Manufacturero	LOGIT	EE.UU.	<u>Muestra:</u> 87,8% <u>Testeo:</u> 87,5%		

**Tabla 1.3. Principales modelos centrados de predicción de insolvencia (Sector Industria)
(Continuación)**

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Yi-Chung Hu y Fang-Mei Tseng (2005)	904 empresas	Industrial Manufacturero	NN	UK	<u>Muestra:</u> 81,64%		
			MDA		<u>Testeo:</u> 83,75%		
			LOGIT		<u>Muestra:</u> 77,94%		
			PROBIT		<u>Testeo:</u> 80,00%		
					<u>Muestra:</u> 78,36%		
Gómez et al. (2008)	364 empresas	Industrial Manufacturero	LOGIT	ESPAÑA	<u>Testeo:</u> 87,50%		
					<u>Testeo:</u> 83%		
					80%		
					80%		
Lin (2009)	254 empresas	Industrial Manufacturero	MDA	TAIWAN	<u>Muestra:</u> 84,3%		
					<u>Testeo:</u> 78,5%		
					74,1%		
					<u>Muestra:</u> 86,4%		
					<u>Testeo:</u> 76,0%		
			LOGIT		<u>Muestra:</u> 89,4%		
					<u>Testeo:</u> 86,4%		
					75,4%		
					<u>Muestra:</u> 82,1%		
					<u>Testeo:</u> 71,7%		
			NN		<u>Muestra:</u> 82,1%		
					<u>Testeo:</u> 86,6%		
					74,3%		
					<u>Muestra:</u> 86,6%		
					<u>Testeo:</u> 75,0%		

**Tabla 1.3. Principales modelos centrados de predicción de insolvencia (Sector Industria)
(Continuación)**

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Appiah y Abor (2009)	62 empresas	Industrial Manufacturero	Z-score	UK	<u>Muestra:</u> 97,3%		
Jae Kwon Bae (2012)	1.888 empresas	Industrial Manufacturero	RSVM	COREA DEL SUR	<u>Muestra:</u> 82,35%		
Callejón et al. (2013)	1.000 empresas	Industrial Manufacturero	NN	EUROPA	<u>Muestra:</u> 92,5% <u>Testeo:</u> 92,1%		
Zhang et al. (2013)	92 empresas	Industrial Manufacturero	SVM LOGIT	CHINA	<u>Testeo:</u> 90,48% <u>Testeo:</u> 95,24%		
Grünenberg y Lukason (2014)	11.600 empresas	Industrial Manufacturero	NN LOGIT	ESTONIA	<u>Quebradas:</u> 84%	57%	62%
					<u>Sanas:</u> 85%	84%	83%
					<u>Quebradas:</u> 72%	43%	38%
					<u>Sanas:</u> 88%	88%	88%

NOTA: Elaboración propia

Tabla 1.4. Principales modelos centrados de predicción de insolvencia (Sector Construcción)

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Abidali y Harris (1995)	31 empresas	Construcción	MODELO Z MODELO A	U.K.	<u>Modelo Z:</u> 90% <u>Modelo A:</u> 50%		
Mínguez Conde (2006)	126 empresas	Construcción	MODELO LOGIT MODELO COX	ESPAÑA	<u>Error tipo I:</u>		
					3,85%		
					<u>Error tipo II:</u>		
					19,26%	7,69%	11,54%
					11,54%	7,69%	7,69%
					<u>Error tipo I:</u>		
					7,69%		
					3,85%		
					15,38%		
					15,38%		
					<u>Error Tipo II:</u>		
					11,54%		
Treewichayapong et al. (2011)	96 empresas	Construcción e inmobiliario	MODELO LOGIT MODELO DE COX	TAILANDIA	<u>Muestra:</u> 94%	94%	96%
					<u>Muestra:</u> 66%	94%	84,4%

**Tabla 1.4. Principales modelos centrados de predicción de insolvencia (sector Construcción)
(Continuación)**

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Gill de Albornoz y Giner (2013)	4.600 empresas	Construcción e inmobiliario	MODELO LOGIT	ESPAÑA	<u>Error total S. Construcción frente</u> <u>Modelo General: 45% vs. 33%</u> <u>Error total S. Inmobiliario frente</u> <u>Modelo General: 41% vs. 53%</u>		

NOTA: Elaboración propia

Tabla 1.5 Principales modelos centrados de predicción de insolvencia (Sector Comercio y Servicios)

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Huo (2006)	324 empresas	Comercio Y Servicios	Modelo de Altman Modelo de Springate Modelo de Fulmer	EE.UU.	Los mejores predictores de insolvencia son el working capital, los beneficios retenidos y el EBIT.		
He y Kamath (2006)	354 empresas	Comercio Y Servicios	Modelo de Ohlson Modelo de Shumway	EE.UU.	<u>Muestra</u> : 80%	65%	70%
					<u>Muestra</u> : 92%	78%	72%
Youn y Zheng (2010)	31 empresas	Comercio Y Servicios	LOGIT NN	EE.UU.	<u>Muestra</u> : 88,1%	76,2%	
					<u>Testeo</u> : 95,0%	85,0%	
					<u>Muestra</u> : 88,1%	78,6%	
					<u>Testeo</u> : 95,0%	80,0%	
Kim (2011)	33 empresas	Comercio Y Servicios	MDA LOGIT SVM	COREA DEL SUR	<u>Muestra</u> : 72,6%		
					<u>Muestra</u> : 80,0%		
					<u>Muestra</u> : 96,0%		

**Tabla 1.5 Principales modelos centrados de predicción de insolvencia (Sector Comercio y Servicios)
(Continuación)**

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Keener (2013)	1.203 empresas	Comercio Y Servicios	LOGIT	EE.UU.	<u>Mejores predictores insolvencia:</u> 1) Número de empleados; 2) Rentabilidad sobre activos (ROA); 3) Margen de cash flow; 4) Ratio Deuda/Capital; 5) Ratio caja/pasivos corrientes.		

NOTA: Elaboración propia

Tabla 1.6 Principales modelos centrados de predicción de insolvencia (Sector Hostelería)

AUTOR Y AÑO	MUESTRA	SECTOR	METODOLOGÍA	PAÍS	RESULTADOS		
					Años antes de quiebra		
					1	2	3
Gu y Gao (2000)	14 empresas (10 restaurantes y 4 hoteles)	Hospitality	MDA	EE.UU.	<u>Muestra:</u> 93%		
Gu (2002)	18 restaurantes	Hospitality	MDA	EE.UU.	<u>Muestra:</u> 92% <u>Testeo:</u> 80%		
Kim y Gu (2006a)	16 empresas (restaurantes, hoteles y casinos)	Hoteles y restaurantes	LOGIT	EE.UU.	<u>Muestra:</u> 91%	84%	
Kim y Gu (2006b)	18 restaurantes	Restaurantes	LOGIT	EE.UU.	<u>Muestra:</u> 94% <u>Testeo:</u> 93%		
Youn y Gu (2010a)	102 empresas hoteleras	Restaurantes	LOGIT NN (MLP)	EE.UU.	<u>Muestra:</u> 95%	85%	
					<u>Muestra:</u> 95%	80%	
Park y Hancer (2012)	40 empresas (hoteles, restaurantes y empresas de servicios)	Hospitality	LOGIT NN (MLP)	EE.UU.	<u>Muestra:</u> 100% <u>Muestra:</u> 100%		

NOTA: Elaboración propia

1.5 FORMULACIÓN DE HIPÓTESIS

Después del análisis efectuado de la literatura sobre predicción de insolvencia, se ha constatado la existencia de un patrón definido por lo que se refiere a la construcción de modelos descentrados frente a modelos centrados, siendo los primeros mucho más numerosos que los segundos. Sin embargo, lo que no ha sido posible es identificar ninguna tendencia, ni a favor ni en contra, en el uso de un tipo de modelo u otro como señala Bellovary et al. (2007). Se deriva, pues, que no es posible inducir una conclusión definitiva sobre la superioridad de un tipo de modelo frente a otro.

Una posible explicación a esta ausencia de un patrón definido de desarrollo en la literatura, en la elaboración de un tipo u otro de modelos, podría basarse en la utilización de muestras heterogéneas de empresas pertenecientes a diferentes sectores o de muestras centradas en empresas que operan en el sector industrial en un sentido amplio. Como afirman Smith y Liou (2007) esta tipología de muestras pudiera ser la razón de la discrepancia entre los buenos resultados de clasificación ex-post y los comparativamente decepcionantes resultados de validación ex-ante, ya que los ratios utilizados como variables independientes pueden verse afectados por efectos sectoriales diferenciales.

En este sentido, es innegable, como afirman Platt y Platt (1990), que los sectores económicos muestran diferencias significativas con respecto a los factores de producción, ciclos de vida de sus productos, estructura competitiva y modos de distribución, lo que da lugar a diferencias importantes en diversas medidas de salud financiera. Hay que apuntar que esta

idea ya fue intuita por Lincoln (1984), Mensah (1984) y McDonald y Morris (1984), quienes analizaron la influencia de los efectos sectoriales en los ratios financieros y su aplicación en la predicción del fracaso empresarial, llegando a la conclusión de que el modelo obtenido para una muestra heterogénea de empresas tiene menos capacidad predictiva que el elaborado para un único sector.

Por su parte, McGur y De-Vaney (1998) llegaron a los mismos resultados cuando aplicaron cinco modelos, desarrollados en otros tantos trabajos, usando empresas de distintos sectores a una única muestra de empresas dedicadas al comercio al por menor.

Asimismo, Bhargava et al. (1998) encontraron pocas similitudes entre los modelos elaborados para una muestra industrial y otra comercial, si bien no se obtuvieron diferencias significativas entre los porcentajes de acierto, lo que indicaba que era posible aplicarlos a una combinación de muestras.

En un intento de superar estas dificultades, algunos autores adoptaron un enfoque denominado "orientado al sector". Este enfoque se basa en el ajuste de los ratios de las empresas por el valor mediano (Izan, 1984) o medio (Platt y Platt, 1990) por el respectivo ratio en el sector de cada empresa en un momento de tiempo. En ambos trabajos, los resultados de clasificación en los modelos, con ratios no ajustados, eran inferiores a los obtenidos con ratios ajustados al sector, confirmando que este enfoque era de utilidad para modelos donde las empresas representan una extensa selección de sectores industriales (Izan, 1984).

Igualmente, Chava y Jarrow (2004) también evaluaron la importancia de incluir los efectos sectoriales en la predicción del fracaso empresarial, analizando diez sectores industriales principales y clasificándolos en cuatro grupos. Incluyeron esta agrupación en los modelos mediante variables dummy, que resultaron ser significativas.

En definitiva, es posible que la ausencia de una conclusión sobre la superioridad de un modelo descentrado sobre un modelo centrado y viceversa sea debida a que no se hayan podido comparar de forma homogénea un tipo de modelo y otro dada a la disparidad de metodologías, enfoques, bases de datos disponibles, periodos temporales y países, entre otras cuestiones.

En consecuencia, este hueco existente en la literatura, que no permite dilucidar la superioridad de los modelos descentrados frente a los modelos centrados, nos ha motivado a formular las siguientes hipótesis:

Hipótesis 1 (H_1): *Las variables predictoras de insolvencia de un modelo global son distintas a las de los modelos sectoriales.*

Hipótesis 2 (H_2): *La introducción de variables cualitativas sectoriales en un modelo global mejora su capacidad de predicción.*

Hipótesis 3 (H_3): *Un modelo global predice la insolvencia correctamente utilizando muestras específicas de los diferentes sectores.*

BIBLIOGRAFÍA CAPITULO 1

Abdullah, N.A.H., Halim, A., Ahmad, H. y R. M. Rus (2008). Predicting Corporate Failure of Malaysia's listed companies: comparing multiple discriminant analysis, logistic regression and the hazard model. *International Research Journal of Finance and Economics*, 15, 201-217.

Abidali, A.F. y F. Harris. (1995). A methodology for predicting company failure in the construction industry. *Construction Management and Economics*, 13, 189-196.

Aharony, J., Jones, Ch. P. y I. Swary. (1980). An analysis of risk and return characteristics of Corporate Bankruptcy using capital market data. *The Journal of Finance*, 35, 4, 1001-1016.

Ahn H. y K. Kim. (2009). Bankruptcy prediction modelling with hybrid case-based reasoning and genetic algorithms approach. *Application Software Computing*, 9, 599-607.

Akers, M. (2007). A Review of Bankruptcy Prediction Studies: 1930 to Present. *Journal of Financial Education*, 33 (Winter), 1-42.

Alam, P., Booth, D., Lee, K. y T. Thordarson (2000). The use of fuzzy clustering algorithm and self-organizing neural networks for identifying potentially failing banks: An experimental study. *Expert Systems with Applications*, 18, 185-199.

Alici, P.Y. (1996). *Neural networks in corporate failure prediction. The UK experience*. Neural Networks in

Financial Engineering. Mostafa, A. Moody, J. y Weigend, A. (eds). Singapore. *World Scientific*, 393-406.

Altman E.I., Marco, G. y F. Varetto (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (The Italian experience). *Journal of Banking and Finance*, 18, 505-529.

Altman, E.I., Y. H. Eom y D.W. Kim (1995). Failure Prediction: Evidence from Korea. *Journal of International Financial Management and Accounting*, 6, 230-249.

Altman, E.I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23, 4, 589-609.

Altman, E.I. (1993). *Corporate Financial Distress and Bankruptcy*. 2nd ed. New York. John Wiley & Sons.

Altman, E.I. y E. Hotchkiss. (2006). *Corporate Financial Distress and Bankruptcy. Predict and avoid Bankruptcy, Analyze and Invest in Distressed Debt*. Third Edition. Wiley Finance.

Altman, E.I., Marco, G. y F. Varetto (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and network. *Journal of Banking and Finance*, 18, 3, 505-529.

Aggarwal, R. y J.W. Goodell (2014). Cross-national differences in access to finance: Influence of culture and

institutional environments. *Research in International Business and Finance*, 31, 193–211.

Anandarajan, M., Lee, P. y A. Anandarajan (2001). Bankruptcy prediction of financially stressed firms: An examination of the predictive accuracy of artificial neural networks. *Intelligent Systems Accounting Finance Management*, 10, 69–81.

Appetiti, A. (1984). Identifying unsound firms in Italy. An attempt to use trend variables. *Journal of Banking and Finance*, 8, 2, 269-279.

Appiah, K.O. y J. Abor. (2009). Predicting corporate failure: some empirical evidence from the UK. *Benchmarking: An International Journal*, 16, 3, 432-444.

Arkaradejdachachai, C. (1993). Study of corporate turnaround. Using a probability of bankruptcy. *Tesis doctoral*, Universidad de Missouri.

Arquero, J.L., Abad, M.C. y S.M. Jiménez (2009). Procesos de fracaso empresarial en Pymes. Identificación y contrastación empírica. *Revista Internacional de la Pequeña y Mediana Empresa*, 2, 1, 64-77.

Atiya, A.F. (2001). Bankruptcy Prediction for Credit Risk Using Neural Network: A Survey and New Results. *IEEE Transactions on Neural Networks*, 12, 4, 929-935.

Aziz, A., Emanuel, D.C. y G.H. Lawson (1988). Bankruptcy Prediction. An investigation of cash flow based models. *Journal of Management Studies*, 25, 5, 419-437.

Aziz, A. y G. Lawson (1989). Cash Flow Reporting and Financial Distress Models: Testing of Hypotheses. *Financial Management*, 18, 1, 55-63.

Back, B., Laitinen, T. y K. Sere (1996). Neural networks and genetic algorithms for bankruptcy predictions. *Expert Systems with Applications*, 11, 407-413.

Baldwin, J. y G.W. Glezen (1992). Bankruptcy prediction using quarterly financial statement data. *Journal of Accounting, Auditing and Finance*, 7, 269-289.

Bartoloni, E. y M. Baussola (2014). Financial Performance in Manufacturing Firms: A comparison between parametric and non-parametric approaches, *Business Economics*, 49, 1, 32-45.

Beaver, W.H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 5, 71-111.

Bell, T., Ribar, G. y J. Verchio (1990). Neural nets versus logistic regression: A comparison of each model's ability to predict commercial bank failures. Proceedings of the 1990 D&T, University of Kansas Symposium on Auditing Problems.

Bellovary, J., Giacomino, D. y M. Akers (2007). A Review of Bankruptcy Prediction Studies: 1930 to Present. *Journal of Financial Education*, 33 (Winter), 1-42.

Bernstein, L.A. (1984). *Análisis de Estados Financieros*. Ed. Deusto. Bilbao.

Betz, F., Oprică, S., Peltonen, T.A. y P. Sarlin (2014). Predicting distress in European banks. *Journal of Banking and Finance*, 45, 225–241.

Bhargava, M., Dubelaar, C. y T. Scott (1998). Predicting bankruptcy in the retail sector: an examination of the validity of key measures of performance. *Journal of Retailing and Consumer Services*, 5, 2, 105-117.

Bongini, P., Ferri, G. y H. Hahm (2000). Corporate bankruptcy in Korea. Only the strong survive?. *The Financial Review*, 35, 4, 31–50.

Boritz, J.E., Kennedy, D.B. y A. Albuquerque (1995). Predicting corporate failure using a neural network approach. *Intelligent Systems in Accounting, Finance and Management*, 4, 95–111.

Boritz, J. y D. Kennedy (1995). Effectiveness of Neural Network Types for Prediction of Business Failure. *Expert Systems with Applications*, 9, 4, 503-512.

Borrajó, M., Barúque, B., Corchado, E., Bajo, J. y J. Corchado (2011). Hybrid neural intelligent system to predict business failure in small-to-medium-size

enterprises. *International Journal Neural Systems*, 21, 4, 277–296.

Bose, I. y R. Pal (2006). Predicting the survival or failure of click-and-mortar corporations: a knowledge discovery approach. *European Journal of Operations Research*, 174, 959–982.

Brigham, E.F. y L.C. Gapenski (1994). *Financial Management: Theory and Practice*. Orlando, FL: The Dryden Press.

Callejón, A.M., Casado, A.M., Fernández, M.A. y J.I. Peláez (2013). A System of Insolvency Prediction for industrial companies using a financial alternative model with neural networks. *International Journal of Computational Intelligence Systems*, 4, 1-13.

Camillo, A.A., Connolly D.J. y W.G. Kim (2008). Success and Failure in Northern California. Critical success factors for independent restaurants. *Cornell Hospitality Quarterly*, 49, 364-380.

Casey, C.J. Jr. (1980). The usefulness of accounting ratios for subjects' predictions of corporate failure: Replication and Extensions. *Journal of Accounting Research*, 18, 2, 603-613.

Casey, C. y N. Bartczak (1985). Using Operating Cash Flow Data to Predict Financial Distress: Some Extensions. *Journal of Accounting Research*, 23, 1, 384-401.

Celentani, M., García-Posada, M. y F. Gómez (2011). *The Spanish Business Bankruptcy Puzzle and the Crisis*. Fundación de Estudios de Economía Aplicada. Documento de Trabajo, nº 2010-2011.

Chang Lee, K., Han, I. y Y. Kwon (1996). Hybrid neural network models for bankruptcy predictions. *Decision Support Systems*, 18, 63-72.

Charalambous, C., Chatitou, A. y F. Kaourou (2000). Comparative analysis of artificial neural network models: Application in bankruptcy prediction. *Annals of Operations Research*, 99, 1, 403-419.

Charitou, A., Neophytou, E. y C. Charalambous (2004). Predicting corporate failure. Empirical evidence for the UK. *European Accounting Review*, 13, 3, 465–497.

Chauhan, N., Ravi, V. y D.K. Chandra (2009). Differential evolution trained wavelet neural networks: Application to bankruptcy prediction in banks. *Expert Systems with Applications*, 36, 4, 7659-7665.

Chava, S. y R.A. Jarrow (2004). Bankruptcy Prediction with Industry Effects. *Review of Finance*, 8, 537-569.

Chen, M.Y. (2011). Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches. *Computers and Mathematics with Applications*, 62, 4514-4524.

Chen, S., Härdle, W.K. y R.A. Moros (2011). Modeling default risk with support vector machines. *Quantitative Finance*, 11, 1, 135-154.

Chen, M.Y. (2011). Predicting corporate financial distress based on integration of decision tree classification and logistic regression. *Expert Systems with Applications*, 38, 11261-1127.

Cheng, C., Chen, C.L. y C.J. Fu (2006). Financial Distress Prediction by a Radial Basis Function Network with Logit Analysis Learning. *Computer and Mathematics with Applications*, 51, 579-588.

Cho, M. (1994). Predicting business failure in the hospitality industry: An application of logit model. *PhD Dissertation*. Virginia Polytechnic Institute and State University.

Cho, S., Hong, H. y B.Ch. Ha (2010). A hybrid approach based on the combination of variable selection using decision trees and case-based reasoning using the Mahalanobis distance: For bankruptcy prediction. *Expert Systems with Applications*, 4, 3482-3488.

Cielen, A., Ludo, P. y K. Vanhoof (2004). Bankruptcy prediction using a data envelopment analysis. *European Journal of Operations Research*, 154, 526–532.

Coats, P.K., y L.F. Fant (1993). Recognizing financial distress patterns using a neural network tool. *Financial Management*, 22, 3, 142-155.

Coats, P.K. y L.F. Fant (1992). A neural network approach to forecasting financial distress. *The Journal of Business Forecasting Methods and Systems*, 10, 4, 9-12.

D'Antoni, J., Mishra, A. y S. Chintawar (2009). *Predicting Financial Stress in Young and Beginning Farmers in the United States*. Department of Agricultural Economics and Agribusiness, Louisiana State University.

Darayseh, M., Waples, E. y D. Tsoukalas (2003). Corporate failure for manufacturing industries using firms specifics and economic environment whit logit analysis. *Managerial Finance*, 29, 8, 23-37.

Deakin, E.B. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research*, 10, 1, 167-179.

De Andrés, J., Landajo, M. y P. Lorca (2005). Forecasting business profitability by using classification techniques. A comparative analysis based on a Spanish case. *European Journal of Operations Research*, 167, 2, 518–542.

De Andrés, J., Landajo, M. y P. Lorca (2012). Bankruptcy prediction models based on multinorm analysis: An alternative to accounting ratios, *Knowledge-Based Systems*, 30, 67-77.

Diamond, J. H. (1976). Pattern recognition and the detection of corporate failure. *PhD. Dissertation*, New York University.

Dietrich, J., Arcelus, F.J. y G. Srinivisan (2005). Predicting Financial Failure: Some Evidence from New Brunswick Agricultural Co-ops. *Annals of Public and Cooperative Economics*, 76, 2, 179-194.

Dietrich, R. (1984). Discussion of the methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* (Supplement 1984).

Dimitras, A., S. Zanakis y C. Zopounidis (1996). A survey of business failures with an emphasis on prediction methods and industrial applications. *European Journal of Operations Research*, 90, 487-513.

Dimitras, A.I., Slowinski, R. y R. Susmaga (1999). Business failure prediction using rough sets. *European Journal of Operations Research*, 114, 263–280.

Ding, Y., Song, X. y Y. Zeng (2008). Forecasting financial condition of Chinese listed companies based on support vector machine. *Expert Systems with Applications*, 34, 3081–3089.

Edminster, R. (1972). An empirical test of financial ratio analysis for small business failure prediction. *Journal of Financial and Quantitative Analysis*, 7, 2, 1477-1493.

El Hennawy, R. y R. Morris (1983). The significance of base year in developing failure prediction models. *Journal of Business Finance and Accounting*, 10, 2, 209-223.

El-Temtamy, O. (1995). Bankruptcy prediction: A comparative study on logit and neural networks. *PhD Dissertation*. Middle Tennessee University.

Espahbodi, P. (1991). Identification of problem banks and binary choice models. *Journal of Banking and Finance*, 15, 1, 53-71.

Etemadi, H., Rostamy, A. y H. Dehkordi (2009). A genetic programming model for bankruptcy prediction: empirical evidence from Iran. *Expert Systems Applications*, 36, 2, 3199–3207.

Evans, J. y A.L. Borders (2014). Strategically surviving bankruptcy during a global financial crisis: The importance of understanding chapter 15. *Journal of Business Research*, 67, 1, 2738–2742.

Friedman, H., Altman, E.I. y D. Kao (1985). Introducing recursive partitioning for financial classification: the case of financial distress. *Journal of Finance*, 40, 269–291.

Fletcher, D. y E. Goss (1993). Forecasting with neural networks. An application using bankruptcy data. *Information and Management*, 24, 159-167.

García, V. y M.A. Fernández (1992). *Solvencia y Rentabilidad de la Empresa Española*. Instituto de Estudios Económicos. Madrid.

Gentry, J., Newbold, P. y D. Whitford (1985). Predicting Bankruptcy: If Cash Flow's Not the Bottom Line, What Is?. *Financial Analysts Journal*, 41, 5, 47-56.

Gesel, T., Baesens, B., Suykens, J., Poel, D., Baestaens, D. y M. Willekens (2006). Bayesian kernel based classification for financial distress detection. *European Journal of Operations Research*, 172, 979-1003.

Gepp, A., Kumar, K. y S. Bhattacharya (2010). Business failure prediction using decision trees. *Journal of Forecasting*, 29, 6, 536-555.

Gilbert, L. R., Menon, K. y K.B. Schwartz (1990). Predicting bankruptcy for firms in financial distress. *Journal of Business Finance and Accounting*, 17, 1, 161-192.

Gill de Albornoz, B. y B. Giner (2013). Predicción del fracaso empresarial en los sectores construcción e inmobiliario: modelos generales versus específicos. *Universia Business Review*, Tercer Trimestre, 118-131.

Gombola, M.J., Haskins, M.E., Ketz, J. E. y D.D. Williams (1987). Cash Flow in Bankruptcy Prediction. *Financial Management*, 16, 4, 55-65.

Gómez, M^a Elena, De la Torre, J.M. y I. Román (2008). Análisis de sensibilidad temporal en los modelos de predicción de insolvencia: una aplicación a las pymes industriales, *Revista Española de Financiación y Contabilidad*, Vol. XXXVII, 137, 85-111.

Grover, J. (2003). Validation of a cash flow model. A non Bankruptcy approach. *Ph.D.Dissertation*, Nova Southeastern University.

Grüenberg, M. y O. Lukason (2014). Predicting Bankruptcy of Manufacturing Firms. *International Journal of Trade, Economics and Finance*, 5, 1, 93-97.

Gu, Z. (2002). Analyzing bankruptcy in the restaurant industry: A multiple discriminant model. *International Journal of Hospitality Management*, 21, 1, 25-42.

Gu, Z. y L. Gao (2000). A multivariate model for predicting business failures of hospitality firms. *Tourism and Hospitality Research. The Survey Quarterly Review*, 2, 1, 37-49.

Guan, Q. (1993). Development of optimal network structures for back-propagation-trained neural networks. *PhD dissertation*. University of Nebraska.

Habilollah, J. y S. Farnoosh (2009). The comparison Artificial Neural Networks and Multi Decimal Analysis Models for Forecasting Bankruptcy and Financial Distress, *Proceedings of the World Congress on Engineering*, vol. II, London.

Hanweck, G. (1977). Predicting bank failures. Research Papers in Banking and Financial Economics, Financial Studies Section, Board of Governors of the Federal Reserve System, Washington D.C.

He, Y. y R. Kamath (2006). Business Failure Prediction in Retail Industry: An empirical evaluation of generic bankruptcy models. *Academy of Accounting and Financial Studies Journal*, 10, 2, 97-110.

Henebry, K. (1996). Do cash flow variables improve the predictive accuracy of a Cox proportional hazards model for bank failure?. *The Quarterly Review of Economics and Finance*, 36, 3, 395-409.

Hing-Ling Lau, Amy (1987). A Five-State Financial Distress Prediction Model. *Journal of Accounting Research*, 25, 1, 127-138.

Hopwood, W., McKeown, J. y J. Mutchler (1989). A Test of the Incremental Explanatory Power of Opinions Qualified for Consistency and Uncertainty. *The Accounting Review*, 44, 1, 28-48.

Hu, Y.C. y F.M. Tseng (2005). Applying Back propagation Neural Network to Bankruptcy Prediction. *International Journal of Electronic Business Management*, 3, 2, 79-103.

Hui, X.-F. y J. Sun (2006). An application of support vector machine to companies financial distress prediction. Lecture Notes. *Artificial Intelligent*, 3885, 274–282.

Huo, Y. (2006). Bankruptcy Situation Model in Small Business: the case of restaurant firms, *Hospitality Review*, 24, 2, 49-58.

Izan, H. (1984). Corporate distress in Australia. *Journal of Banking and Finance*, 8, 303-320.

Jae Kwon Bae (2012). Predicting financial distress of the South Korean manufacturing industries, *Expert Systems with Applications*, 39, 9159-9165.

Jo, H. y I. Han (1996). Integration of case-based forecasting, neural network, and discriminant analysis for bankruptcy prediction. *Expert Systems with Applications*, 11, 415–422.

Jo, H. y I. Han (1997). Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis. *Expert Systems with Applications*, 13, 97–108.

Johnsen, T. y R.W. Melicher (1994). Predicting Corporate Bankruptcy and Financial Distress: Information Value Added by Multinomial Logit Models. *Journal of Economics and Business*, 46, 269-286.

Jones, S. y D.A. Hensher (2007). Forecasting Corporate Bankruptcy: Optimizing the Performance of the Mixed Logit Model. *Abacus*, 43, 241–264.

Karels G.V. y Arun J. Prakash (1987). Multivariate Normality and Forecasting of Business Bankruptcy. *Journal of Business Finance and Accounting*, 14, 4, 573-593.

Kahya, E. y P. Theodossiou (1999). Predicting Corporate Financial Distress: A time series CUSUM Methodology.

Review of Quantitative Finance and Accounting, 13, 323-345.

Keener, M. (2013). Predicting the financial failure of retail companies in the United States, *Journal of Business and Economics Research*, 11, 8, 373-380.

Ketz, J.E. (1978). The effect of general price-level adjustments on the predictive ability of financial ratios. *Journal of Accounting Research*, 16, 273-284.

Kim, H y Z. Gu (2006 a). A logistic regression analysis for predicting bankruptcy in the Hospitality Industry. *The Journal of Hospitality Financial Management*, 14, 1, 17-34.

Kim, H y Z. Gu (2006 b). Predicting Restaurant Bankruptcy. A Logit Model in Comparison with a Discriminant Model. *Journal of Hospitality and Tourism Research*, 30, 4, 474-493.

Kim, Soo Y. (2011). Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression and multivariate analysis, *The Service Industries Journal*, 31, 3, 441-468.

Koh, H. y L.N. Killough (1990). The use of multiple discriminant analysis in the assessment of the going-concern status of an audit client. *Journal of Business Finance and Accounting*, 17, 2, 179-192.

Korol, T. (2013). Early warning models against bankruptcy risk for Central European and Latin American enterprises. *Economic Modelling*, 31, 22–30.

Kotler, P., Bowen, J. y J. Makens (1996). *Marketing for Hospitality & Tourism*. Upper Saddle River, Editorial Prentice Hall.

Kwak, W., Shi, Y. y G. Kou (2011). Bankruptcy prediction for Korean firms after the 1997 financial crisis: using a multiple criteria linear programming data mining approach. *Review of Quantitative Finance and Accounting*, 38, 4, 441-453.

Kwak, W., Yong Shi y K. Gang (2012). Bankruptcy prediction for Korean firms after the 1997 financial crisis: using a multiple criteria linear programming data mining approach. *Review of Quantitative Finance and Accounting*, 38, 441-453.

Lamothe, P., Gómez, P. y A. Partal (2010). *Utilidad de los métodos cuantitativos (ratings) en la valoración del riesgo de crédito (insolvencia) en las pymes*. En Reflexiones para la Reforma Concursal (Coord. Ariza, M. J.). Ed. Reus, 9-40.

Lane, W., Looney, S. y J. Wansley (1986). An application of the Cox proportional hazards model to bank failure. *Journal of Banking and Finance*, 10, 511-531.

Langford, D., Iyagba, R. y D.M. Komba (1993). Prediction of solvency in construction companies. *Construction Management and Economics*, 11, 317-325.

Lau, A. (1987). A five-state financial distress prediction model. *Journal of Accounting Research*, 25, 1, 127-138.

Lee, K.C., Han, I. y Y. Kwon (1996). Hybrid Neural Network models for bankruptcy predictions. *Decision Support Systems*, 18, 63-72.

Lee, K. (2001). Pattern classifications and clustering algorithms with supervised and unsupervised neural networks in financial applications. *PhD Dissertation*. Kent State University.

Lee, S., Koh, Y. y K.H. Hang (2011). Moderating effect of capital intensity on the relationship between leverage and financial distress in the US restaurant industry. *International Journal of Hospitality Management*, 30, 429-438.

Leshno, M. y Y. Spector (1996). Neural Network Prediction Analysis: The bankruptcy case. *Neurocomputing*, 10, 125-147.

Levitan, A. y J. Knoblett (1985). Indicators of Exceptions to the Going Concern Assumption. *Auditing: A Journal of Practice and Theory*, 5, 1, 26-39.

Li, H., Huang, H., Sun, J. y C. Lin (2010). On sensitivity of case-based reasoning to optimal feature subsets in business failure prediction. *Expert System with Applications*, 37, 7, 4811-4821.

Libby, R. (1975). Accounting ratios and the prediction of failure: Some behavioral evidence. *Journal of Accounting Research*, 13, 1, 150-161.

Li, H., Sun, J. y J. Wu (2010). Predicting business failure using classification and regression tree: an empirical comparison with popular classical statistical methods and top classification mining methods. *Expert Systems with Applications*, 37, 8, 5895–5904.

Lin, T.H. (2009). A cross model study of corporate financial distress prediction in Taiwan: Multiple Discriminant Analysis, logit, probit and neural networks models. *Neurocomputing*, 72, 3507-3516.

Lin, F., Yeh, C.C. y M.Y. Lee (2011). The use of hybrid manifold learning and support vector machines in the prediction of business failure. *Knowledge-Based Systems*, 24, 95–101.

Lincoln, M. (1984). An empirical study of the usefulness of accounting ratios to describe levels of insolvency risk. *Journal of Banking and Finance*, 8, 321-340.

Lo, A.W. (1986). Logit versus discriminant analysis: a specification test and application to corporate bankruptcies. *Journal of Econometrics*, 31, 151-178.

Logue, L. y J. Merville (1972). Financial policy and market expectations. *Financial Management*, 1, 3, 37-44.

Lussier, R.N. y S. Pfeiffer (2000). A comparison of business success versus failure variables between U.S. and central eastern Europe Croatian entrepreneurs. *Entrepreneurship Theory and Practice*, 24, 4, 59–67.

Maghyereh, A.I. y B. Awartani (2014). Bank distress prediction: Empirical evidence from the Gulf Cooperation Council countries. *Research in International Business and Finance*, 30, 126–147.

Malcolm, S. y D. Liou (2007). Industrial sector and financial distress. *Managerial Auditing Journal*, 22, 4, 376–391.

Martin, D. (1977). Early warning of bank failures: A logit regression approach. *Journal of Banking and Finance*, 1, 249–276.

Martin del Brio, B. and C. Serrano-Cinca (1995). Self-organizing neural networks: The financial state of Spanish companies. Article in *Neural Network in the Capital Markets*, Refenes (ed.). Chichester: Wiley, 341–357.

Mateos-Ronco, A., Marín-Sánchez, M^a del Mar, Marí-Vidal, Sergio y E. Seguí-Mas (2011). Los modelos de predicción del fracaso empresarial y su aplicabilidad en cooperativas agrarias. *CIRIEC-España, Revista de Economía Pública, Social y Cooperativa*, 70, abril, 179–208.

McDonald, B. y M.H. Morris (1984). The statistical validity of the ratio method in financial analysis: An empirical

examination. *Journal of Business Finance and Accounting*, 11, 1, 89-97.

McGurr, P.T. y S.A. DeVaney (1998). Predicting business failure of retail firms: An analysis using mixed industry models. *Journal of Business Research*, 43, 169-176.

McKee, T. (1995). Predicting Bankruptcy via Induction. *Journal of Information Technology*, 10, 26-36.

McKee, T.E. y M. Greenstein (2000). Predicting bankruptcy using recursive partitioning and a realistically proportioned data set. *Journal of Forecasting*, 19, 219-230.

McKee, T.E. (2000). Developing a bankruptcy prediction model via rough sets theory. *Intelligent Systems Accounting Finance Management*, 9, 159-173.

Mensah Y.M. (1983). The Differential Bankruptcy Predictive Ability of Specific Price Level Adjustments: Some Empirical Evidence. *The Accounting Review*, 53, 2, 228-246.

Mensah, Y.M. (1984). An examination of the stationarity of multivariate bankruptcy prediction models: A methodological study. *Journal of Accounting Research*, 22, 1, 380-395.

Messier, Jr.,W. y J. Hansen (1988). Inducing rules for expert system development: An example using default and bankruptcy data. *Management Science*, 34, 12, 1403-1415.

Meyer, P. y H. Pifer (1970). Prediction of bank failures. *Journal of Finance*, 25, 4, 853-868.

Min, J.H. y Y.C. Lee (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28, 128-134.

Mínguez Conde, J.L. (2006). El fracaso empresarial en la empresa constructora. *Tesis Doctoral*. Universidad de Valladolid.

Moyer, R.Ch. (1977). Forecasting financial failure: A re-examination. *Financial Management*, 6, 1, 11-17.

Mulford, C.W. y E.E. Comiskey (2005). *Creative cash-flow reporting. Uncovering Sustainable Financial Performance*. John Wiley and Sons. New Jersey, 316 y ss.

Muller, G.H., Steyn-Bruwer y W.D. Hamman (2009). Predicting financial distress of companies listed on the JSE. A comparison of techniques. *South African Journal of Business Management*, 40, 1, 21-32.

Nasir, M.L., John, R.I., Bennett, S.C., Russell, D.M. y A. Patel (2000). Predicting corporate bankruptcy using artificial neural networks. *Journal of Applied Accounting Research*, 5, 3, 30-51.

Neophytou, E. y C.M. Molinero (2004). Predicting corporate failure in the UK: A multidimensional scaling approach.

Journal of Business Finance and Accounting, 31, 5/6, 677-710.

Norton, C.L. y R.E. Smith (1979). A comparison of general price level and historical cost financial statements in the prediction of bankruptcy. *The Accounting Review*, 54, 1, 72-87.

Nour, M. (1994). Improved clustering and classification algorithms for the Kohonen self-organizing neural network. *PhD Dissertation*. Kent State University.

Odom, M.D. y R. Sharda (1990). A neural networks for bankruptcy prediction. *IEEE INNS International Joint Conference on Neural Networks*, 2, 17-21, 163-168

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18, 1, 109-131.

Olsen, M., Bellas, C. y L.V. Kish (1983). Improving the prediction of restaurant failure through ratio analysis. *International Journal of Hospitality Management*, 2, 187-193.

Pantalone, C. y M. Platt. (1987a). Predicting commercial bank failure since deregulation. Federal Reserve Bank of Boston New England Economic Review (July/ August), 37-47.

Pantalone, C. y M. Platt. (1987b). Predicting failure of savings and loan associations. *AREUEA Journal*, 15, 2, 46-64.

Park, S. y M. Hancer (2012). A comparative study of logit and artificial neural networks in predicting bankruptcy in the hospitality industry. *Tourism Economics*, 18, 2, 311-338.

Park, C. S. y I. Han (2002). A case-based reasoning with the feature weights derived by analytic hierarchy process for bankruptcy prediction. *Expert Systems with Applications*, 23, 255–264.

Parsa H.G., Self J., Sydnr-Busso S. y H.J. Yoon (2011). Why Restaurants Fail? Part II. The Impact of Affiliation, Location, and Size on Restaurant Failures: results from a survival analysis. *Journal of Food Service Business Research*, 14, 360-370.

Patterson, D. (2001). Bankruptcy prediction: A model for the casino industry. *PhD Dissertation*. University of Nevada-Las Vegas.

Pendharkar, P.C. (2002). A Potential Use of Data Envelopment Analysis for the Inverse Classification Problem. *Omega*, 30, 3, 243-248.

Pettway, R. y J. Sinkey, Jr.. (1980). Establishing on-site banking examination priorities: An early warning system using accounting and market information. *Journal of*

Finance, 35,1, 137-150.

Platt, H.D. y M.B. Platt (1990). Development of a class of stable predictive variables: The case of bankruptcy prediction. *Journal of Business Finance and Accounting*, 17,1, 31-51.

Platt, H.D. y M.B. Platt (1991). A note in the use of industry-relative ratios in bankruptcy prediction. *Journal of Banking and Finance*, 15, 6, 1183-1194.

Platt, H.D. y M.B. Platt (2008). Financial distress comparison across three global regions. *Journal of Risk and Financial Management*, 1, 1, 129-162.

Pindado, J. Rodrigues, L. y Ch. De la Torre (2008). Estimating Financial Distress Likelihood. *Journal of Business Research*, 61, 995-1003.

Parsa, H. G., Self, J. T., Njite, D. y T. King (2005). Why restaurants fail?. *Cornell Hotel and Restaurant Administration Quarterly*, 46, 304-22.

Premachandra, I.M., Chen, Y. y J. Watson (2011). DEA as a tool for predicting corporate failure and success: a case of bankruptcy assessment. *Omega*, 39, 6, 620–626.

Rafiei, F.M., Manzari, S.M. y S. Bostanian (2011). Financial health prediction models using artificial neural networks, genetic algorithm and multivariate discriminant analysis:

iranian evidence. *Expert Systems with Applications*, 38, 10210–10217.

Ravisankar, P., Ravi, V. y I. Bose (2010). Failure prediction of dotcom companies using neural network-genetic programming hybrids. *Informatic Science*, 180, 1257–1267.

Ravisankar, P. y V. Ravi (2010). Financial distress prediction in banks using Group Method of Data Handling neural network, counter propagation neural network and fuzzy ARTMAP. *Knowledge-Based Systems*, 23, 823–831.

Rose, P. y J. Kolari (1985). Early warning systems as a monitoring device for bank condition. *Quarterly Journal of Business and Economics*, 24,1, 43-60.

Ryu, Y. y W. Yue (2005). Firm bankruptcy prediction: experimental comparison of isotonic separation and other classification approaches. *IEEE Trans. Syst. Man Cybern–Part A: Systems Humans*, 21, 4, 265–276.

Salchenberger, L., Cinar, E. y N. Lash (1992). Neural networks: A new tool for predicting bank failures. *Decision Sciences*, 23, 899-916.

Santomero, A. y J. Vinso (1977). Estimating the probability of failure for comercial banks and the banking system. *Journal of Banking and Finance*, 1, 2, 185-205.

Sarkar, S. y R.S. Sriram (2001). Bayesian models for early warning of bank failures. *Management Science*, 47, 11, 1457–1475.

Scherrer, P.S. y T.J. Mathison (1996). Investment strategies for REIT inventories. *Real State Review*, 26, 1, 5-10.

Shah, J. y M. Murtaza (2000). A neural network model based clustering procedure for bankruptcy prediction. *American Business Review*, 18, 2, 80-86.

Sangjae Lee y Wu Sing Choi (2013). A multi-industry bankruptcy prediction model using back propagation neural network and multivariate discriminant analysis. *Expert Systems with Applications*, 40, 2941-2946.

Sean, S., Hyun, S. y W. Kim (2011). Dimensions of brand equity in the Chain Restaurant Industry. *Cornell Hospitality Quarterly*, 52, 429-437.

Serrano-Cinca, C. (1996): Self Organizing neural networks for financial diagnosis. *Decision Support Systems*, 17, 3, 227-238.

Serrano Cinca, C. (1997). Feedforward neural networks in the classification of financial information. *European Journal of Finance*, 1997, 3, 3, 183-202.

Serrano-Cinca, C. y B. Gutiérrez-Nieto (2013). Partial Least Square Discriminant Analysis for bankruptcy prediction. *Decision Support Systems*, 3, 1245-1255.

Shin, K-S, y Y.J. Lee (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications*, 23, 3, 321-328.

Shin, K. S., Lee, T. S. y H.J. Kim (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, 28, 127–135.

Shuk-Wern Ong, Voon Choong Yap y R.W.L. Khong (2011). Corporate failure prediction: a study of public listed companies in Malaysia. *Managerial Finance*, 37, 6, 553-564.

Shumway, T. (2001): Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business*, 74, 1, 101-124.

Sinkey, Jr., J. (1975). A multivariate statistical analysis of the characteristics of problematic banks. *Journal of Finance*, 30, 1, 21-36.

Skogsviki, K. (1980). Current cost accounting ratios as predictors of business failure. The Swedish case. *Journal of Business, Finance and Accounting*, 17, 1, 137-160.

Smith, M. Y D.K. Liou (2007). Industrial sector and financial distress. *Managerial Auditing Journal*, 22, 4, 376-391.

Stroe, R. y N. Barbuta-Misu (2010). Predicting the financial performance of the building sector enterprises-Case Study

of Galati County (Romania), *The Review of Finance and Banking*, 2, 1, 29-39.

Soo Cheong J., Park, K. y E. Lee-Ji (2012). Estimating Cash flow-at-Risk (CFaR). A Comparables Approach for Restaurant Firms. *Cornell Hospitality Quarterly*, 52, 232-240.

Spicka, J. (2013). The financial condition of the construction companies before bankruptcy. *European Journal of Business and Management*, 5, 23, 160-169.

Susicky, J. (2011). Applicability of Bankruptcy Models at Agricultural Companies. *Acta VSFS*, 3/2011, 5, 241-252.

Sun, T., Chang, N. y G. Lee (1999). Dynamics of modeling in data mining. Interpretive approach to bankruptcy prediction. *Journal of Management Information Systems*, 16, 1, 63-85.

Sun, J. y H. Li (2008). Listed companies financial distress prediction based on weighted majority voting combination of multiple classifiers. *Expert Systems with Applications*, 35, 818–827.

Sun, L. y P. Shenoy (2007). Using Bayesian networks for bankruptcy prediction: Some methodological issues. *European Journal of Operations Research*, 180, 738–753.

Sun, J., Li, H., Huang, Q-H. y K.Y. He (2014). Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and

featuring approaches. *Knowledge-Based Systems*, 57, 41–56.

Takahashi, K., Kurokawa, Y. y K. Watase (1984). Corporate bankruptcy prediction in Japan. *Journal of Banking and Finance*, 8, 2, 229-247.

Taffler, R. (1984). Empirical models for the monitoring of UK corporations. *Journal of Banking and Finance*, 8, 2, 199-227.

Tam, K. (1991). Neural network models and the prediction of bank bankruptcy. *Omega*, 19, 5, 429–445.

Tam, K. y M. Kiang (1992). Managerial applications of neural networks: the case of bank failure prediction. *Management Science*, 38, 7, 926–947.

Theodossiou, P. (1991). Alternative models for assessing the financial condition of Business in Greece. *Journal of Business and Accounting*, 18, 5, 697-720.

Theodossiou, P. (1993). Predicting Shifts in the Mean of a Multivariate Time Series Process: An Application in Predicting Business Failures. *Journal of the American Statistical Association*, 88, 422, 441-449.

Tisshaw, H.J. (1976). Evaluation of downside risk using financial ratios, *M.Sc. thesis*. City University Business School. London.

Treewichayapong, S., Chunhachinda, P. y C. Padungsaksawasdi (2011). Bankruptcy Prediction of Real

Estate Firms in Thailand, *The International Journal of Finance*, 23, 1, 6672-6691.

Tsai, C. (2009). Feature selection in bankruptcy prediction. *Knowledge-Based Systems*, 22, 2, 120–127.

Tsai, C., Hsu, Y. y D.C. Yen (2014). A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing*, 24, 977–984.

Tseng, F., y Y. Hu (2010). Comparing four bankruptcy prediction models: Logit, quadratic, interval logit, neural and fuzzy neural networks. *Expert Systems with Applications*, 37, 1846-1853.

Tsukuda, J. y S. Baba (1994). Predicting Japanese Corporate Bankruptcy in terms of financial date using Neural Network. *Computers and Industrial Engineering*, 27, 445-448.

Tzong-Huei Lin (2009): A cross model study of corporate financial distress prediction in Taiwan: Multiple discriminant analysis, logit, probit and neural networks models. *Neurocomputing*, 72, 3507-3516.

Varetto, F. (1998). Genetic algorithms applications in the analysis of insolvency risk. *Journal of Banking and Finance*, 22, 1421–1439.

Vavrina, J., Hampel, D. y J. Janová (2013). New approaches for the financial distress classification in

agribusiness. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, vol. LXI, 4, 1177-1182.

Wang, B. (2004). Strategy changes and internet firm survival. *PhD. Dissertation*. University of Minnesota.

Wang, Y., Wang, S. y K.K. Lai (2005). A new fuzzy support vector machine to evaluate credit risk. *IEEE Transactions on Fuzzy Systems*, 13, 6, 820–831.

Ward, T. (1994). An empirical study of the incremental predictive ability of Beaver's naive Operating Cash Flow measure using four-state ordinal models of financial distress. *Journal of Business Finance and Accounting*, 21, 4, 547-561.

Wasilewski, M. y M. Madra (2008). An assessment of the agricultural enterprises' solvency with the usage of the Altman Model. *Acta Oeconomica et Informatica*, 2, 50-55.

Wilcox, J. W. (1973). A prediction of business failure using accounting data. *Journal of Accounting Research*, 11, 163-179.

Wilson R.L. y R. Sharda (1994). Bankruptcy prediction using neural networks. *Decision Support Systems*, 11, 545-557.

Wu, C.H., Tzeng, G.H., Goo, Y.J. y W.C. Fang (2007). A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy. *Expert System with Applications*, 32, 397-408.

Wu, D., Liang, L. y Z. Yang (2008). Analysing the financial distress of Chinese public companies using probabilistic neural networks and multivariate discriminate analysis. *Socio-Economic Planning Science*, 42, 3, 206–220.

Wu, W. (2011). Improving classification accuracy and causal knowledge for better credit decisions. *International Journal of Neural Systems*, 21, 4, 297–309.

Xu, M. y C. Zhang (2009). Bankruptcy prediction: the case of Japanese listed companies, *Review of Accounting Studies*, 14, 534-558.

Yang, Z. R., Platt, M. B. y H.D. Platt (1999). Probabilistic neural networks in bankruptcy prediction. *Journal of Business Research*, 44, 67–74.

Yeh, C. C., Chi, D. J. y M.F. Hsu (2010). A hybrid approach of DEA, rough set and support vector machines for business failure prediction. *Expert Systems with Applications*, 37, 1535–1541.

Yi-Chung Hu y Fang-Mei Tseng (2005). Applying back propagation neural networks to bankruptcy prediction, *International Journal of Electronic Business Management*, 3, 2, 97-103.

Youn, H, y Z. Gu (2010a). Predict US restaurant firm failures: the artificial neural network mode versus logistic regression model. *Tourism and Hospitality Research*, 10, 3, 171-187.

Youn, H, y Z. Gu (2010b). Predicting Korean lodging firm failures: An artificial neural network model along with a logistic regression model. *International Journal of Hospitality Management*, 29, 120-127.

Zacharakis, A., Meyer, D. y J. De Castro (1999). Differing perceptions of new venture failure: A matched exploratory study of venture capitalists and entrepreneurs. *Journal of Small Business Management*, 37, 1-14.

Zanganeh, T., Rabiee, M. y M. Zarei (2011). Applying Neuro-Fuzzy Model for Bankruptcy Prediction. *International Journal of Computer Applications*, 20, 3, 15-21.

Zavgren, C.V. (1985). Assessing the vulnerability to failure of American industrial firms. *Journal of Business and Accounting*, 12, 1, 19-45.

Zhang, G., Hu, M.Y., Patuwo, B.E. y D.C. Indro (1999). Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operations Research*, 116, 1, 16-33.

Zhang, L., Zhang, L., Teng, W. y Ch. Yibing (2013). Based on Information Fusion Technique with Data Mining in the Application of Finance Early-Warning. *Procedia Computer Science*, 17, 695-703.

Zhou, L. (2013). Performance of corporate bankruptcy prediction models on imbalanced dataset: The effect of sampling methods. *Knowledge-Based Systems*, 41, 16-25.

Zimmer, I. (1980). A lens study of the prediction of corporate failure by bank loan officers. *Journal of Accounting Research*, 18, 2, 629-636.

Zmijewski, M.E. (1984). Methodological Issues Related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22, 59-81.

CAPITULO 2:

MÉTODOS DE CLASIFICACIÓN BINARIOS

En esta tesis pretendemos clarificar si los modelos de predicción de insolvencia específicos para diferentes sectores económicos ofrecen mejores resultados que la aplicación de un modelo global a diferentes muestras formadas por empresas de sectores económicos concretos.

Para lograr este objetivo se estimarán modelos Logit que permitirán analizar la incidencia de diversas variables en la predicción de la insolvencia en los sectores económicos tratados en este estudio.

El análisis multivariante puede llevarse a cabo mediante diversos métodos estadísticos que relacionan una variable explicativa, o dependiente, con otra u otras variables explicativas, que se denominan independientes. Para el estudio que nos ocupa, este análisis permite clasificar las empresas en solventes e insolventes según sean los valores que tomen las diversas variables explicativas.

En general, pueden diferenciarse dentro del análisis multivariante las técnicas descriptivas y las técnicas explicativas. En las técnicas descriptivas, al no existir hipótesis previas, el objetivo es descubrir nuevos conocimientos en la materia objeto de estudio. Habitualmente la finalidad es la reducción de datos o la identificación de grupos homogéneos. En cambio, las técnicas explicativas tienen como objetivo predecir el comportamiento de los sujetos a partir de los valores que toman una serie de variables independientes. En este caso sí se adoptan hipótesis previas que son objeto de

contrastación. El Análisis Discriminante y la Regresión Logística son dos de las técnicas de análisis multivariante más utilizadas.

2.1 ANÁLISIS DISCRIMINANTE

El Análisis Discriminante (AD) es una técnica estadística que permite analizar si existen diferencias entre dos grupos con respecto a determinadas variables y en qué sentido se dan dichas diferencias. Asimismo, es una herramienta que permite elaborar un modelo sistemático de clasificación de individuos desconocidos en alguno de los grupos analizados.

En este sentido, además de la capacidad predictiva o de clasificación del modelo estadístico, también contaría con una utilidad explicativa o descriptiva, en tanto que este método permitiría determinar qué variables independientes contribuyen en mayor medida a la clasificación en uno u otro grupo, pudiendo diferenciarse el AD predictivo del AD explicativo.

Este método estadístico requiere del cumplimiento de una serie de supuestos para su correcta aplicación. Sanchis (2000) establecía como principales requisitos los siguientes:

- Las variables que describen a cada uno de los miembros de las observaciones de cada grupo han de estar distribuidas como una distribución normal multivariante dentro de cada grupo.

- Las matrices de covarianza de los grupos deben ser iguales en todos ellos.

- Los grupos han de ser discretos, mutuamente excluyentes e identificables.

- Las muestras deben ser aleatorias dentro de cada grupo y, por lo tanto, representativas de la población de la que provienen.

Asimismo habría que resaltar la importancia de otros supuestos, como es el caso de la ausencia de la multicolinealidad para la obtención de un modelo de AD correcto (Sánchez, 1984).

Dependiendo del número de categorías que pueda adoptar la variable dependiente, es posible diferenciar dos tipos: Análisis Discriminante Simple, en cuyo caso la variable dependiente podrá adoptar tan solo dos valores, o el Análisis Discriminante Múltiple (MDA), en el caso de que la variable dependiente pueda adoptar un número de categorías superior a dos. Igualmente, el número de funciones discriminantes vendrá determinada por el número de categorías, menos 1, que tengamos en el fenómeno analizado.

Conviene recordar que el Análisis Discriminante surgió del trabajo de Fisher (1936) ⁶. Supone la combinación lineal de dos o más variables independientes que permite establecer una diferencia entre los grupos definidos a priori. La ponderación de cada variable se lleva a cabo de tal forma que maximice la varianza entre-grupos frente a la varianza intra-grupos. Así, la ecuación de la función discriminante vendrá determinada por la siguiente expresión:

$$Z_{jk} = a + W_1X_{1k} + W_2X_{2k} + \dots + W_nX_{nk} \quad (2.1)$$

Donde:

Z_{jk} = puntuación Z discriminante de la función discriminante j para el objeto k .

a = constante.

W_i = ponderación discriminante para la variable independiente i .

X_{ik} = variable independiente i para el objeto k .

El procedimiento utilizado para estimar los distintos pesos (W_j) sería el de Mínimos Cuadrados Ordinarios

⁶Ronald Aylmer Fisher (1892-1962): científico británico inventor del análisis discriminante, del método de máxima verosimilitud y del diseño estadístico de experimentos. Trabajó en Rothamsted, una estación experimental agrícola en Inglaterra, y fue profesor de Eugenética en la Universidad de Londres. Además de sus numerosas contribuciones a todas las ramas de Estadística, que le sitúan como el padre de esta disciplina en el siglo XX, fue célebre como genetista, investigador agrario y biólogo (Peña, 2002).

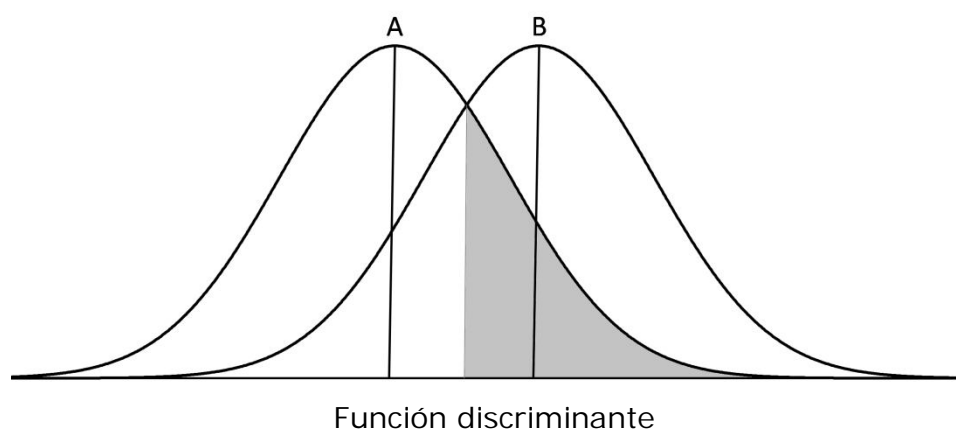
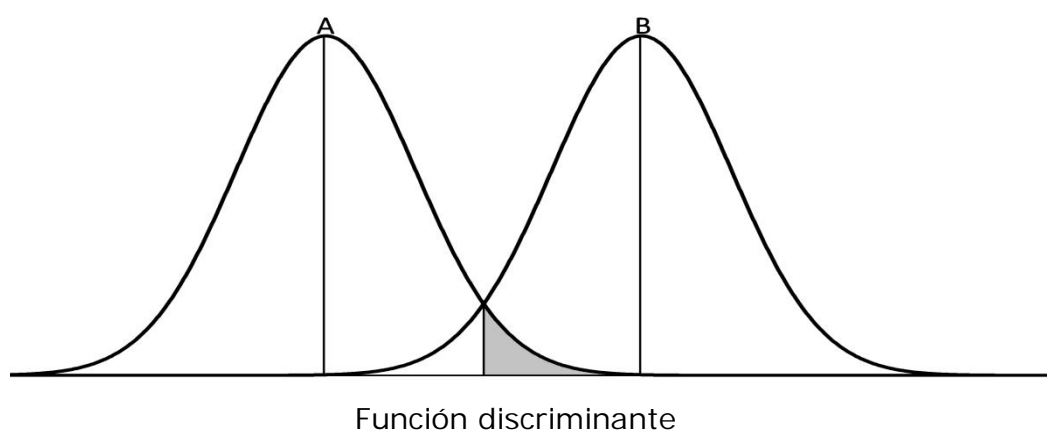
(MCO), que estima el valor de los parámetros que minimizan el término de error, consiguiendo así unas predicciones más certeras.

Hair et al. (1999) establecieron que el análisis discriminante es la técnica estadística apropiada para contrastar la hipótesis de que las medias de un conjunto de variables independientes para dos o más grupos son iguales. Para llevarlo a cabo, el análisis discriminante multiplica cada variable independiente por su correspondiente ponderación y, posteriormente, suma estos productos. El resultado es una única puntuación Z discriminante compuesta para cada individuo en el análisis. Promediando las puntuaciones discriminantes para todos los individuos dentro de un grupo particular, obtenemos la media del grupo, a la que se conoce con el nombre de centroide. Cuando el análisis comprende dos grupos, existen dos centroides; con tres grupos, hay tres centroides, y así sucesivamente. Los centroides indican la posición más común de cualquier individuo perteneciente a un determinado grupo. Así, una comparación de los centroides de los distintos grupos revelaría los apartados en los que se encuentran los grupos, en relación con la dimensión que se está contrastando.

Igualmente, el contraste para la significación estadística de la función discriminante es una medida generalizada de la distancia entre los centroides de los grupos. Se calcula comparando las distribuciones de las puntuaciones discriminantes para dichos grupos. Si el solapamiento en la distribución es pequeño implica que la

función discriminante separa bien los grupos. Si el solapamiento es grande, la función es un mal discriminador entre los grupos (Hair et al. 1999).

GRÁFICO 2.1 EJEMPLOS DE FUNCIONES DISCRIMINANTES



Fuente: Hair et al. (1999)

En el gráfico 2.1 se representan dos distribuciones discriminantes para dos funciones distintas. El área sombreada se correspondería con aquellos casos en los que se clasifica erróneamente objetos del grupo A en el B. En el primero de los diagramas, el área sombreada es menor y representa la distribución de puntuaciones discriminantes para una función que separa bien los grupos, mientras que en el diagrama de abajo se muestra la distribución de puntuaciones discriminantes de una función que corresponde a un discriminador relativamente malo entre los grupos A y B.

Una vez estimada la función discriminante, el siguiente paso a realizar se correspondería con la comprobación del nivel de ajuste global del modelo estimado. En este sentido cabe resaltar que el modelo AD no establece, propiamente, una clasificación de los individuos en las categorías establecidas, debiendo para ello construirse las denominadas matrices de clasificación. Previamente se calcula la puntuación discriminante de cada observación, para posteriormente evaluar las diferencias de grupo sobre estas puntuaciones discriminantes. Una vez obtenida esta matriz de clasificación, es posible valorar la precisión en la predicción.

Hay que puntualizar que el cálculo de la puntuación discriminante se realiza para cada grupo en la propia observación, catalogándose posteriormente en aquel grupo en el que haya obtenido una mayor puntuación, lo cual se

conseguirá determinando un punto de corte que fijará dónde clasificar cada objeto.

Una vez obtenidas las puntuaciones discriminantes y derivado el punto de corte, se podrá construir la matriz de clasificación, la cual mostrará las observaciones correctamente clasificadas y, conociéndose, por tanto, la capacidad predictiva del modelo.

Obtenida la función discriminante, y valorada la capacidad predictiva del modelo, ya se estaría en condiciones de analizar la importancia o relevancia de cada una de las variables incluidas en el modelo globalmente, con objeto de obtener conclusiones. Este es el objetivo de lo que anteriormente se ha descrito como Análisis Discriminante Explicativo.

2.2 MODELOS DE RESPUESTA CUALITATIVA

Los modelos de respuesta cualitativa permiten relacionar una variable dependiente cualitativa, con una o más variables independientes, las cuales podrán ser dicotómicas, politómicas o continuas. Un modelo de respuesta cualitativa permitirá predecir la probabilidad de que una respuesta sea elegida.

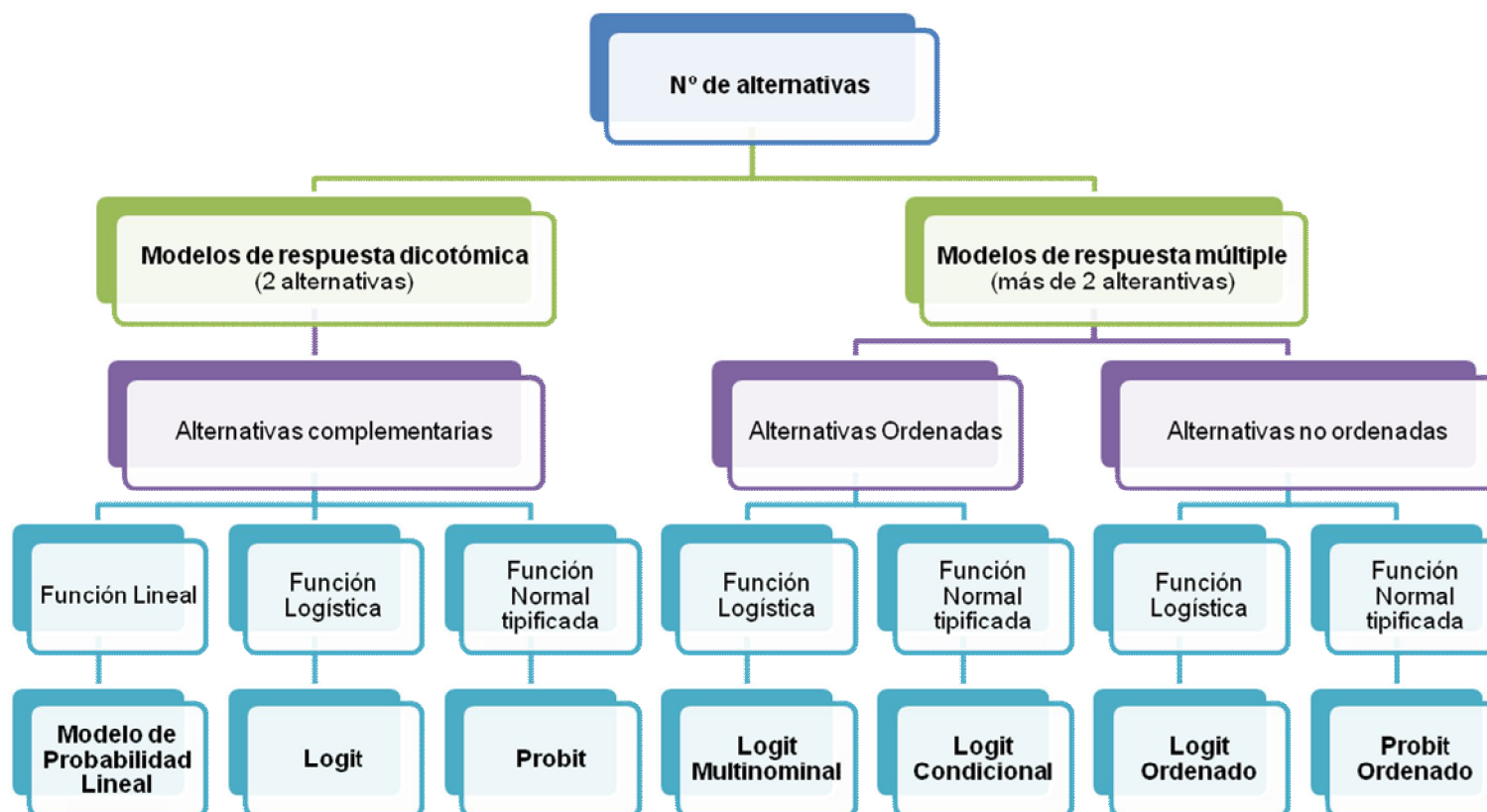
El modelo de regresión es, sin duda, una de las principales herramientas utilizadas en Econometría y constituye el primer paso para interpretar los resultados derivados de las aplicaciones empíricas. Permite contrastar

hipótesis teóricas sobre las relaciones entre variables económicas (Levy y Varela, 2003).

Según el número de alternativas incluidas en la variable endógena, se distinguen los modelos de respuesta dicotómica frente a los denominados modelos de respuesta o elección múltiple. Según la función utilizada para la estimación de la probabilidad, se distingue entre el modelo de probabilidad lineal truncado, el modelo Logit y el modelo Probit. Según que las alternativas de la variable endógena sean excluyentes o incorporen información ordinal, se diferencia entre los modelos con datos no ordenados y los modelos con datos ordenados. Dentro de los primeros, según que los regresores hagan referencia a aspectos específicos de la muestra o de las alternativas entre las que se pretende elegir, se distingue entre los modelos multinominales y los condicionales (Medina, 2003).

En el gráfico 2.2 se puede apreciar una clasificación general de los distintos modelos de respuesta cualitativa.

GRÁFICO 2.2 CLASIFICACIÓN DE LOS MODELOS DE RESPUESTA CUALITATIVA



Fuente: Medina (2003)

2.2.1 MODELO LINEAL DE PROBABILIDAD

La expresión matemática de un modelo lineal de probabilidad es la siguiente:

$$Y_i = a + \beta_1 X_{1j} + \beta_2 X_{2j} + \dots + \beta_n X_{nj} + \varepsilon_i \quad (2.2)$$

Donde:

$$Y_i = \begin{cases} 1 & \text{si ocurre una alternativa} \\ 0 & \text{en caso contrario} \end{cases}$$

X_{nj} = variables explicativas.

β_n = coeficientes de regresión de cada una de las variables explicativas.

ε_i = variable aleatoria que se distribuye como una normal $N(0, \delta^2)$.

El Modelo Lineal de Probabilidad se puede interpretar en términos probabilísticos, en el sentido de que un punto o valor concreto de la recta de regresión nos indicaría la probabilidad de que ocurriera el acontecimiento estudiado. Es decir, Y_i se puede considerar como la estimación de la probabilidad de que ocurra el acontecimiento objeto de estudio ($Y_i = 1$) siguiendo el criterio: valores próximos a cero se corresponden con una baja probabilidad de ocurrencia del acontecimiento analizado (menor cuanto más próximos a cero); mientras que a valores próximos a

1 se les asigna una probabilidad elevada de ocurrencia (mayor cuanto más próximos a 1) (Medina, 2003). Igualmente, Medina (2003) señala que las principales limitaciones de la estimación del modelo por Mínimos Cuadrados Ordinarios (MCO), serían las siguientes:

- El valor estimado de Y_i podría no estar comprendido en el rango entre 0 y 1, lo cual carecería de lógica al interpretarse el resultado como una probabilidad.

- La perturbación aleatoria (ε_i) puede no seguir una distribución normal, al tratarse de una distribución binaria o dicotómica. Si bien el incumplimiento de la normalidad no invalida la estimación por MCO, la ausencia de normalidad imposibilita el uso de los estadísticos habituales utilizados para realizar el contraste de hipótesis, tales como la t-Student y la F-Snedecor, en la medida en que estos estadísticos requerirían el cumplimiento de la hipótesis de normalidad en la perturbación aleatoria.

- Problemas de heterocedasticidad. En caso de que la perturbación aleatoria fuera heterocedástica y la estimación del modelo mediante MCO obtuviera unos estimadores de los coeficientes de regresión con varianza no mínima, estos estimadores serían no eficientes.

- El coeficiente de determinación R^2 está subestimado. La suma de los cuadrados de los residuos ($\sum e_i^2$) es mayor de lo habitual debido a la forma específica en que se distribuye la nube de puntos de una variable

dicotómica. Dado que el cálculo del coeficiente de determinación R^2 se ve afectado por Σe_i^2 , el R^2 calculado en la estimación por MCO es más pequeño de lo que realmente debería ser.

Son precisamente los inconvenientes presentados por los modelos probabilísticos lineales, tanto en lo que se refiere a estimación como a interpretación, los que han llevado a la búsqueda de otros modelos alternativos que permitieran estimaciones más fiables de las variables dicotómicas.

2.2.2 MODELOS DE PROBABILIDAD NO LINEAL

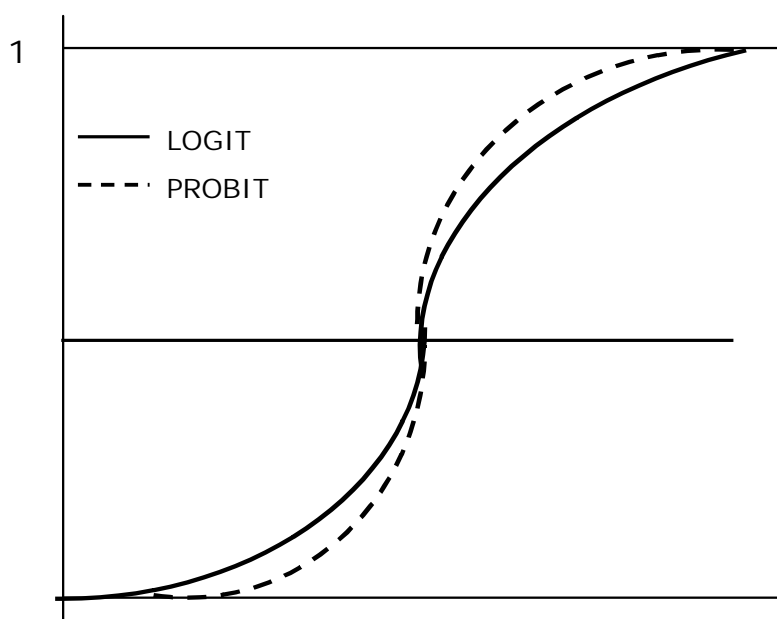
Los modelos de probabilidad no lineal permiten salvar la principal limitación de los modelos lineales, por cuanto la función de especificación utilizada garantiza un resultado en la estimación comprendido en el rango entre 0 y 1. Las funciones de distribución cumplirían estos requisitos, en tanto son funciones continuas que toman valores entre 0 y 1.

Entre las funciones de distribución más habituales se encuentra la función logística, que ha dado lugar a los modelos Logit, y la función de distribución de la normal tipificada, que ha dado lugar a los modelos Probit.

Existe una gran similitud entre las curvas de los modelos Logit y Probit, lo cual se puede apreciar en el gráfico 2.3.

En el caso de los modelos Probit, la función de distribución tan solo puede obtenerse a partir de una función integral, dado que es una distribución de la normal tipificada. Esto presenta una mayor complejidad en el cálculo e interpretación de los parámetros obtenidos con respecto a los modelos logísticos, lo que unido a la similitud en los resultados, ha potenciado la utilización de los modelos Logit, que serán analizados en mayor profundidad en el siguiente apartado.

GRÁFICO 2.3 REPRESENTACIÓN GRÁFICA DE LOS MODELOS LOGIT Y PROBIT



Fuente: Elaboración propia

2.2.2.1 Regresión Logística

Si bien el Análisis Discriminante (AD) es apropiado cuando la variable dependiente es una variable no métrica,

pero cuando esta variable es de tipo binario es aconsejable emplear, para resolver un problema de clasificación, la regresión logística (Logit) debido a las limitaciones que presenta AD. En este sentido, Logit no requeriría cumplir con supuestos tan estrictos y resultaría ser mucho más robusta cuando dichos supuestos no se cumplen, circunstancia muy habitual cuando se tratan datos económicos.

Sin embargo, no conviene olvidar que Logit no deja de ser una técnica de clasificación, en la que la variable dependiente considera exclusivamente dos categorías, si bien parte de unos supuestos menos restrictivos y permite que el modelo incorpore variables cualitativas (Visauta, 2003).

Además, Logit es un modelo no lineal, a pesar de que contiene una combinación lineal de parámetros y observaciones de las variables explicativas. La función logística se encuentra acotada entre 0 y 1, proporcionando, por tanto, la probabilidad de que un elemento se encuentre en uno de los dos grupos establecidos.

Es decir, a partir de un suceso dicotómico, Logit predice la probabilidad de que el evento tenga o no lugar. Si la estimación de la probabilidad es superior a 0,5 entonces la predicción es que sí pertenece a ese grupo, y en caso contrario, supondría que pertenece al otro grupo considerado.

Para estimar el modelo se parte del cociente entre la probabilidad de que un suceso ocurra y la probabilidad de que este no ocurra.

La probabilidad de que un suceso ocurra vendrá determinada por la expresión:

$$P(Y_i = 1/x_i) = \frac{e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}} \quad (2.3)$$

donde β_0 es el término constante del modelo y los β_1, \dots, β_k los coeficientes de las variables.

Así, el ratio entre las dos probabilidades (lo que se conoce como el *Odds ratio*) sería:

$$\begin{aligned} Odds &= \frac{P(Y_i = 1)}{1 - P(Y_i = 1)} = \frac{1/(1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)})}{1/(1 + e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)})} = \\ &= \frac{1 + e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}} = e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)} \end{aligned} \quad (2.4)$$

Los coeficientes estimados $(\beta_0, \beta_1, \dots, \beta_k)$ representan, en realidad, medidas de los cambios en el ratio de probabilidades (Odds ratios). En este sentido, un coeficiente positivo aumenta la probabilidad de ocurrencia del suceso, mientras que un valor negativo disminuye la probabilidad de ocurrencia del mismo (Hair et al., 1999). Los Odds ratios se pueden interpretar como el número de veces que es más probable que ocurra el fenómeno frente a que no ocurra.

Por tanto, si aplicáramos logaritmos, se obtendría una expresión lineal del modelo Y_i^* , en la forma:

$$Y_i^* = \ln \frac{P(Y_i = 1)}{1 - P(Y_i = 1)} = \ln e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)} = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k \quad (2.5)$$

En donde los coeficientes del modelo $(\beta_0, \beta_1, \dots, \beta_k)$ se estimarían aplicando el método de máxima verosimilitud. Este procedimiento implica una serie de pasos: en primer lugar, se especificaría la función de máxima verosimilitud del modelo, que expresaría la probabilidad conjunta para las observaciones independientes consideradas; en segundo lugar, se calcularía el logaritmo neperiano de la función de verosimilitud; en tercer lugar, se llevaría a cabo el cálculo de la derivada del logaritmo neperiano de dicha función respecto a los parámetros que se desean estimar, y, en último lugar, se obtendrían las soluciones al sistema de k ecuaciones planteado o estimadores verosímiles (Levy y Varela, 2003).

Como se puede constatar, si bien la regresión del modelo lineal de probabilidad estima los coeficientes minimizando la desviación de los mínimos cuadrados, la regresión logística maximiza la verosimilitud de que un acontecimiento tenga lugar.

En los modelos Logit, a la hora de valorar la significación individual de cada una de las variables utilizadas en el modelo, se suele utilizar el estadístico de Wald.

Otra cuestión es la de evaluar conjuntamente los resultados. Para ello, debe realizarse una matriz de clasificación en la cual se comparan los valores observados para la variable dependiente con los valores predichos por el modelo, lo que permite obtener la probabilidad de acierto general del modelo estimado. De la misma forma, existen distintos estadísticos que permiten valorar la robustez de la estimación, como ocurre con el test de Hosmer-Lemeshow.

Como se puede deducir de los modelos analizados previamente, Logit resulta una opción idónea para resolver el problema de la clasificación de individuos entre grupos, tanto por la ausencia de limitaciones previas de difícil cumplimiento como por la facilidad de su interpretación, permitiendo no solo obtener un modelo de clasificación, sino interpretar además la relevancia e incidencia de las distintas variables utilizadas para explicar el hecho analizado.

2.2.2.2 Criterios para la selección de modelos estadísticos

La construcción de modelos estadísticos surge de la necesidad de explicar y predecir un fenómeno real que depende de variables.

Generalmente, cuando se trata de modelar este comportamiento, y se plantean varios modelos alternativos, surge el problema de qué modelo seleccionar dentro del conjunto de modelos que se hayan estimado.

La selección de un mejor modelo de la colección de modelos construidos es, a menudo, una elección entre teorías en competencia y se basa en resultados empíricos obtenidos a partir de datos de la muestra.

En este capítulo trataremos los dos métodos empleados en este trabajo de investigación para la selección de modelos de regresión: el método de validación cruzada y los criterios basados en la información estadística.

2.2.2.2.1 Método de Validación Cruzada

La validación cruzada es uno de los métodos más utilizados para evaluar la precisión predictiva de un modelo –validación de resultados-, que se da a priori o desarrollado por un procedimiento de modelización (Allen, 1974; Stone, 1974; Geisser, 1975).

Principalmente, se basa en la división de los datos, donde parte de estos datos se utiliza para la construcción de cada modelo, y el resto de los datos se utiliza para medir la precisión de los modelos de predicción mediante la evaluación de los errores de validación, seleccionándose el modelo con el mejor rendimiento global (Zhang y Yan, 2015).

En este asunto, la validación cruzada (CV) se ha utilizado ampliamente en la minería de datos para la selección del modelo o bien como selección del procedimiento de modelización (Hastie et al., 2009).

Un tema fundamental en la aplicación de CV para la discriminación de modelos es la selección del ratio de división de datos o el tamaño del conjunto de validación. En el presente trabajo, se realiza una división de 70% de la muestra para el conjunto muestra de entrenamiento y un 30% para la muestra de validación.

De los diferentes tipos de CV (validación cruzada dejando uno fuera, de retención y de k -iteraciones), la variante de k -iteraciones es una de las más utilizadas (Cheng y Pecht, 2012). En el método de k -iteraciones, el conjunto de datos se divide en k subconjuntos de igual tamaño. A continuación, se lleva a cabo un proceso iterativo. En cada iteración, se selecciona un subconjunto diferente como datos de validación para estimar la precisión del modelo, y los restantes $k-1$ subconjuntos se utilizan como prueba de entrenamiento (para entrenar el modelo). Para la elección más común de 10 iteraciones (10-CV), nueve subconjuntos se utilizan como pruebas de entrenamiento y el subconjunto restante para la prueba de testeo. Este procedimiento se repite hasta que cada uno de los diez subconjuntos se prueba. El 10-CV realiza, de manera independiente, diez diferentes divisiones del conjunto de datos.

2.2.2.2.2 Criterios basados en la información estadística

Los procedimientos basados en contrastes de hipótesis no penalizan el exceso de parámetros debido a que frecuentemente se selecciona un modelo saturado

como referencia. En cambio, los criterios de selección de modelos basados en la Teoría de la Información equilibran el ajuste del modelo y su complejidad.

Estos criterios toman la forma de una función de verosimilitud penalizada, que es el valor negativo de la del logaritmo de la función de verosimilitud más un término de penalización, que aumenta conforme lo hace el número de parámetros.

Entre los diversos métodos o criterios de selección de modelos basados en la Teoría de la Información, pueden distinguirse tres criterios relevantes dentro de los denominados informacionales. Estos criterios son: el criterio de información de Akaike (1973), el bayesiano de Schwarz (1978) y el de Hannan-Quinn (1979).

2.2.2.2.2.1 El criterio de Akaike (AIC)

El criterio básico de entre los que se fundamentan en la información estadística es el criterio de información de Akaike, que fue inicialmente desarrollado en series temporales.

En el caso general, el criterio de Akaike viene expresado como:

$$AIC = 2k - 2\ln(L) \quad (2.6)$$

donde:

k = número de parámetros.

L = Máximo valor de la función de verosimilitud del modelo estimado.

La idea básica subyacente al uso del criterio de Akaike para la selección de modelos es la maximización del logaritmo de la función de verosimilitud esperada de un modelo determinado.

El segundo término de la expresión (2.6) puede ser interpretado como una medida de la bondad de ajuste, en tanto que el primer término es una penalización creciente conforme aumenta el número de parámetros.

Como señala Takane (1987), el criterio de Akaike enfatiza la bondad del modelo, si bien no pretende identificar al modelo verdadero. Que un modelo sea el que mejor se ajusta a los datos no quiere decir que el modelo será real o verdadero. Más bien significa que el modelo es el mejor de entre los modelos candidatos, en el sentido de que proporciona la aproximación más cercana a la realidad o al verdadero modelo.

El modelo que mejor se ajusta a los datos, podría cambiar en función del tamaño muestral, dado que, con un tamaño muestral mayor, los parámetros de un modelo pueden ser estimados de manera más fiable.

Como puede verse en (2.6) el término de penalización, $2k$, no depende del tamaño muestral de la población considerada. Esto conduce al hecho de que un mismo número de parámetros puede ser seleccionado

mediante el criterio de Akaike, tanto para muestras pequeñas como para muestras grandes.

2.2.2.2.2 El criterio de Schwarz (BIC)

Schwarz (1978) sugirió que el criterio de Akaike podría no ser asintóticamente justificable y presentó un criterio de información alternativo a partir de un enfoque bayesiano. Con este criterio se penaliza el número de parámetros con $\ln(n)$ en lugar de con 2. Así, la expresión del criterio de Schwarz quedaría como sigue:

$$BIC = -2\ln(L) + \ln(n) \times k \quad (2.7)$$

donde:

k = número de parámetros.

L = máximo valor de la función de verosimilitud del modelo estimado.

n = número de observaciones

De esta forma, este criterio pondera la desviación del modelo, medida por $-2\ln(L)$, con el número de parámetros. Si introducimos más parámetros en el modelo mejorará el ajuste, con lo que disminuirá la desviación y este efecto será compensado por el aumento del número de parámetros.

Este criterio, a diferencia del criterio de Akaike, considera el número de observaciones en el término de

penalización, resultando menos favorable a la inclusión de parámetros.

Otra diferencia entre los criterios de Akaike y de Schwarz es que son además los comúnmente más usados, por su distinto objetivo (Peña, 2002). El criterio BIC trata de seleccionar el modelo correcto, con máxima probabilidad a posteriori y es un criterio consistente. El criterio AIC no pretende seleccionar el modelo correcto, ya que admite que el modelo verdadero puede no estar entre los estimados, y trata de obtener el modelo que proporcione mejores predicciones entre los existentes.

Puede demostrarse que, suponiendo que el modelo verdadero puede aproximarse arbitrariamente bien con los estimados al aumentar el tamaño muestral, el criterio AIC es eficiente, escogiendo el modelo que proporciona en promedio mejores predicciones. Sin embargo, en muestras pequeñas o medianas, el criterio AIC tiende a seleccionar modelos con más parámetros de los necesarios (Peña, 2002).

En resumen, el criterio BIC asume que el modelo real está incluido en el conjunto de modelos candidatos y mide el grado de certeza de que un cierto modelo sea el verdadero modelo generador de los datos.

2.2.2.2.2.3 El criterio de Hannan-Quinn (HQC)

El criterio de Hannan y Quinn puede ser considerado una variante del criterio de Schwarz, con una pequeña penalización de la magnitud del tamaño muestral. Hannan y Quinn (1979) sugirieron este criterio de información, inicialmente propuesto para seleccionar el orden de una autoregresión. Así:

$$HQC = -2Ln(L) + 2Ln[Ln(n)] \times k \quad (2.8)$$

donde:

k = número de parámetros.

L = máximo valor de la función de verosimilitud del modelo estimado.

n = número de observaciones.

Al igual que para los criterios AIC y BIC, se selecciona el modelo que minimice el valor de HQC.

BIBLIOGRAFÍA CAPÍTULO 2

Akaike, H. (1973). *Information theory and an extension of the maximum likelihood principle*. In B. N. Petrov and F. Csaki (Eds.), *Second international symposium on information theory*, 267-281.

Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, 19, 6, 716–723.

Allen, D.M. (1974). The relationship between variable selection and data augmentation and a method for prediction. *Technometrics*, 16, 125–127.

Cheng, S. y M. Pecht (2012). Using cross-validation for model parameter selection of sequential probability ratio test. *Expert Systems with Applications*, 39, 8467–8473.

Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of American Statistical Association*, 70, 320–328.

Hair, J. (2008). *Análisis Multivariante*; Editorial Prentice Hall.

Hair, J.F., Anderson R.E., Tatham, R.L. y W.C. Black (1999). *Análisis multivariante*, 5ª edición. Editorial Prentice Hall. Madrid.

Hannan, E.J. y B.G. Quinn (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society, Series B*, 41, 2, 190-195.

Hastie, T., Tibshirani, R. y J. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference,*

and Prediction. Springer, New York.

Kolmogorov, A.N. (1957). On the representations of continuous functions of many variables by superpositions of continuous functions of one variable and addition. *Doklady Akademii Nauk USSR*, 114,5, 953-956.

Levy, J.P. y J. Varela (2003). *Análisis Multivariable para las Ciencias Sociales*. Ed. Prentice Hall. Madrid.

Medina, E. (2003). *Modelización de variables discretas*. Universidad Autónoma de Madrid. Departamento de Economía Aplicada.

Peña, D. (2002). *Análisis de datos multivariantes*. Ed. Mc Graw Hill. Madrid.

Sánchez, J.J. (1984). *Introducción a las técnicas de análisis multivariante aplicadas a las Ciencias Sociales*. Centro de Investigaciones Sociológicas, Madrid.

Sanchis, A. (2000). *Una aplicación del análisis discriminante a la previsión de la insolvencia en las empresas españolas de seguros no-vida*. Tesis doctoral. Universidad Complutense de Madrid. Facultad de Ciencias Económicas y Empresariales.

Schwarz, G. (1978). Estimating the dimension of a model, *Annals of Statistics*, 6, 461-464.

Stone, M. (1974). Cross-validation choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B*, 36, 111-147.

Takane, Y. (1987). Introduction to special section. *Psychometrika*, 52, 3, 316-316.

Varela, J. y A. Rial (2008). *Estadística práctica para la investigación en Ciencias de la Salud*. Netbiblo.

Visauta, B. (2003). *Análisis Estadístico con SPSS para Windows, Vol. II*. E. McGraw-Hill.

Zhang, Y. y Yang, Y. (2015). Cross-validation for selecting a model selection procedure. *Journal of Econometrics*, 187, 95–112.

PARTE II:

INVESTIGACIÓN EMPIRICA

CAPÍTULO 3:

DATOS Y VARIABLES

En el presente capítulo se comenzarán a desarrollar los aspectos metodológicos y empíricos de la investigación.

Con este objetivo, en primer lugar se expondrán las características de la base de datos utilizada para la elaboración de este estudio. Dicha base se ha obtenido a partir de datos contables y corporativos de empresas españolas.

Se han obtenido dos muestras que han sido utilizadas para analizar la predicción de insolvencia 1 y 2 años antes de su declaración en cinco sectores económicos (agricultura, industria, construcción, comercio y servicios, y hostelería) así como en una muestra global compuesta por empresas de los diferentes sectores económicos considerados.

En segundo lugar, se analizarán las variables que en los trabajos de investigación previos han sido consideradas como posibles predictoras de la insolvencia empresarial, su descripción y su forma de cálculo, a partir de la información contable y corporativa obtenida de la base de datos utilizada.

En el capítulo siguiente, se expondrán los resultados obtenidos, realizándose en primer lugar un análisis exploratorio de los datos que componen las dos muestras consideradas, incluyendo un análisis descriptivo de las variables, diversos test univariantes y un análisis de correlación.

Finalmente, para finalizar la investigación empírica se realizará un análisis confirmatorio multivariante, donde además de estimar los modelos estadísticos considerados (Logit), validaremos los mismos utilizando distintas muestras de testeo. Adicionalmente, se analizará la robustez de las estimaciones y predicciones obtenidas mediante tres criterios de selección de modelos como son el criterio de Akaike, el criterio bayesiano de Schwartz y el criterio de Hannan-Quinn.

3.1 DATOS

En el ámbito del presente estudio se considera que una empresa es insolvente si cuenta con el estatus legal de situación concursal, según las consideraciones realizadas por la ley concursal 22/2003 de 9 de julio, así como las siguientes modificaciones realizadas de la misma (Real Decreto Ley 3/2009 de 27 de marzo de medidas urgentes ante la evolución de la situación económica y la Ley 38/2011, de 10 de Octubre).

Conforme a la legislación señalada, se define el concurso de acreedores como aquel procedimiento legal que se origina cuando una persona física o jurídica deviene en una situación de insolvencia en la que no puede hacer frente a la totalidad de los pagos que adeuda. En la Ley 22/2003 se recoge que el actual concurso de acreedores abarca la situación de suspensión de pagos y de quiebra, considerándose la suspensión de pagos como una insolvencia transitoria y la de quiebra como una

insolvencia definitiva, ambas referidas a personas jurídicas⁷.

Además, y como se indicó con anterioridad, en la última modificación realizada de la Ley 38/2011, se da la posibilidad de la situación de pre-concurso, en la cual el deudor puede comunicar formalmente al Juzgado el inicio de las negociaciones con los acreedores para llegar a un acuerdo de financiación, pudiéndose homologar este acuerdo judicialmente en los tres meses desde la fecha de su presentación. En este estudio no se considera la situación de pre-concurso como una situación de insolvencia.

Una vez definido el ámbito de actuación objeto de estudio, profundizaremos en el proceso de obtención de los datos y las características de cada una de las muestras analizadas.

3.1.1 OBTENCIÓN DE LA BASE DE DATOS

La base de datos utilizada en el presente trabajo está formada tanto por empresas solventes como insolventes pertenecientes a cinco sectores económicos diferenciados: agricultura, industria, construcción, comercio y servicios, y sector de hostelería. La identificación de las empresas pertenecientes a cada sector de actividad se ha realizado

⁷ Para análisis del marco normativo de la Unión Europea en materia de solvencia véase González, Perea y Sánchez (2003).

en base la clasificación realizada por los códigos CNAE-2009.

Así, para el sector agricultura se ha considerado el CNAE correspondiente a los códigos primarios siguientes: 01 - Agricultura, ganadería, caza y servicios relacionados con las mismas, y 02 - Silvicultura y explotación forestal.

Para el sector industria se ha considerado el CNAE correspondiente a los códigos primarios siguientes: 10 - Industria de la alimentación, 11 - Fabricación de bebidas, 12 - Industria del tabaco, 13 - Industria textil, 14 - Confección de prendas de vestir, 15 - Industria del cuero y del calzado, 16 - Industria de la madera y del corcho, excepto muebles; cestería y espartería, 17 - Industria del papel, 18 - Artes gráficas y reproducción de soportes grabados, 19 - Coquerías y refino de petróleo, 20 - Industria química, 21 - Fabricación de productos farmacéuticos, 22 - Fabricación de productos de caucho y plásticos, 23 - Fabricación de otros productos minerales no metálicos, 24 - Metalurgia; fabricación de productos de hierro, acero y ferroaleaciones, 25 - Fabricación de productos metálicos, excepto maquinaria y equipo, 26 - Fabricación de productos informáticos, electrónicos y ópticos, 27 - Fabricación de material y equipo eléctrico, 28 - Fabricación de maquinaria y equipo n.c.o.p., 29 - Fabricación de vehículos de motor, remolques y semirremolques, 30 - Fabricación de otro material de transporte, 31 - Fabricación de muebles, 32 - Otras industrias manufactureras.

Para el sector construcción se ha considerado el CNAE correspondiente al código primario 41 - Construcción de edificios.

Para el sector comercio y servicios se ha considerado el CNAE correspondiente a los códigos primarios siguientes: 45 - Venta y reparación de vehículos de motor y motocicletas, 46 - Comercio al por mayor e intermediarios del comercio, excepto de vehículos de motor y motocicletas, 47 - Comercio al por menor, excepto de vehículos de motor y motocicletas, 49 - Transporte terrestre y por tubería, 50 - Transporte marítimo y por vías navegables interiores, 51 - Transporte aéreo, 52 - Almacenamiento y actividades anexas al transporte, 53 - Actividades postales y de correos, 61 - Telecomunicaciones, 62 - Programación, consultoría y otras actividades relacionadas con la informática.

Finalmente, para el sector de hostelería, se ha considerado el CNAE correspondiente a los códigos primarios 55 - Servicios de alojamiento y 56 - Servicios de comidas y bebidas.

Los datos utilizados para este estudio se corresponden al periodo comprendido entre el año 2010 y el año 2012.

Para realizar una investigación sobre la predicción de insolvencia que permita obtener conclusiones extrapolables y que cuente con la robustez metodológica necesaria, se han obtenido datos de una muestra suficientemente

grande como para considerarse representativa de las empresas insolventes que pertenecen a cada sector de actividad objeto de análisis. En este sentido se ha calculado el número de empresas insolventes necesarias, representativas de una población finita a partir de la siguiente expresión:

$$n = \frac{N * Z_a^2 * p * q}{\epsilon^2 * (N-1) + Z_a^2 * p * q} \quad (3.1)$$

donde:

n = número de empresas insolventes representativas del total de la población.

N = total de la población.

Z_a^2 = coeficiente relativo al nivel de confianza o seguridad considerado. Para una confianza del 95% el valor de Z_a será de 1,96.

p = proporción esperada. En caso de no conocerse la proporción, como norma general se utiliza una proporción del 50% ($p=0,5$).

$q = 1 - p$.

ϵ = término de error.

El número de empresas que integran la población de compañías concursadas en España se ha obtenido de la estadística concursal del Instituto Nacional de Estadística (INE), organismo que controla e informa de las situaciones concursales desde el año 2004.

En conjunto, en este estudio, se ha considerado un total de 6.141 empresas en situación de quiebra. El número de empresas seleccionadas aleatoriamente, entre las que han incurrido en procedimientos concursales en España en los distintos sectores económicos considerados, ha sido obtenido para asegurar una muestra suficientemente significativa.

Una vez seleccionado el número mínimo de empresas insolventes, y en consecuencia, de solventes, que deberán componer la muestra de cada sector, se ha procedido a obtener la información económica y financiera necesaria para la labor de investigación, y que se fundamentará en los datos contables de las distintas empresas. Dichos datos se han obtenido de la base de datos SABI (Sistema de Análisis de Balances Ibéricos), la cual se define como una base de datos económica-financiera que incluye más de 1.250.000 empresas españolas y más de 400.000 empresas portuguesas. La información facilitada por dicha base de datos para cada uno de los registros serían los siguientes: información de identificación y contacto, descripción de la actividad, códigos nacionales e internacionales de actividades, forma jurídica, fecha de constitución, número de empleados, consejo de administración, auditores, bancos, accionistas, filiales, participaciones, cotización en bolsa, balances, cuentas de pérdidas y beneficios, ratios, indicadores y tasas de variación. Los datos disponibles en la base de datos SABI se actualizan periódicamente y se obtienen de

fuentes oficiales, tal es el caso del Registro Mercantil o el BORME.⁸

Para los modelos contruidos con información de 1 año antes de la declaración del concurso de acreedores, se han considerado un total de 1.287 empresas en situación de quiebra. Por sectores, en agricultura se han considerado 90 empresas en esta situación, en industria 300, en construcción 300, en comercio y servicios 300 y en el sector de hostelería 297 empresas.

Para los modelos contruidos con información de 2 años antes de la declaración del concurso de acreedores, se han considerado un total de 1.290 empresas en situación de quiebra. Por sectores, en agricultura se han considerado 93 empresas en esta situación, en industria 300, en construcción 300, en comercio y servicios 300 y en el sector de hostelería 297 empresas.

En el estudio a realizar, y según se detallará en el apartado siguiente, se considerarán el mismo número de empresas solventes que insolventes, lo cual ha sido el criterio más comúnmente empleado en la literatura.

Con objeto de validar los modelos a estimar y comprobar su capacidad predictiva, se utilizan asimismo muestras de validación y de testeo, diferentes a las utilizadas en la estimación de los modelos. En

⁸ BORME: Boletín Oficial del Registro Mercantil.

consecuencia se ha procedido a dividir las muestras en dos submuestras diferenciadas: un 70% se utiliza para el entrenamiento del modelo y el 30% restante se utiliza como muestra de validación.

En último lugar, se procede a calcular las diferentes variables explicativas a partir de la información financiera y corporativa de las distintas empresas que compondrán las muestras (obtenidas de la literatura previa analizada).

3.1.2 OBTENCIÓN DE LAS MUESTRAS

Este estudio ha considerado dos muestras distintas con objeto de construir modelos de predicción de insolvencia utilizando información correspondiente a 1 y 2 años antes de la quiebra.

En las dos muestras obtenidas, se ha considerado el mismo número de empresas solventes que de empresas en situación de quiebra que, como se ha expuesto, es un criterio aplicado en gran parte de los estudios de predicción de insolvencia. Zmijewski (1984) examinó la influencia de una muestra no aleatoria en la determinación de los resultados empíricos, y concluyeron, que en general, la utilización de muestras emparejadas por un criterio determinado no afecta a la capacidad predictiva obtenida en el modelo y en la resultante de testar el mismo en la muestra externa. De cualquier forma, más allá de la existencia de sesgos en la selección de las muestras, en los estudios de predicción de insolvencia se han venido utilizando generalmente muestras

emparejadas, siendo al menos el mismo número de empresas las quebradas y las no quebradas.

Una vez expuestas las características generales de las muestras, se detalla la composición de cada una de ellas. La muestra denominada M.I, utilizada para el análisis de la predicción de insolvencia con información de 1 año antes de dicho evento, está compuesta por un total de 2.575 empresas, de las cuales 1.287 empresas son insolventes y 1.288 son solventes.

Igualmente, la muestra utilizada para estimar los modelos de predicción de insolvencia 2 años antes de la quiebra (M.II) está formada también por un total de 2.578 empresas (1.290 empresas insolventes y 1.288 empresas solventes).

El gráfico 3.1 detalla la ficha técnica de muestreo, considerando las dos muestras mencionadas.

GRÁFICO 3.1. FICHA TÉCNICA DEL MUESTREO**PREDICCIÓN DE INSOLVENCIA EN LOS SECTORES
ECONÓMICOS: UN ANÁLISIS COMPARATIVO**

**TESIS DOCTORAL. UNIVERSIDAD DE MÁLAGA. NOVIEMBRE
2015**

Ámbito: España

Universo: Empresas españolas solventes e insolventes pertenecientes a los sectores de agricultura, industria, construcción, comercio y servicios, y hostelería. Periodo 2010-2012.

Tamaño de la muestra: Dos muestras:

- Muestra I: 1.288 empresas solventes y 1.287 insolventes, siendo los datos contables de las empresas insolventes un año antes de la entrada en quiebra.
- Muestra II: 1.288 empresas solventes y 1.290 insolventes, siendo los datos contables de las empresas insolventes dos años antes de la entrada en quiebra.

Procedimiento de Muestreo: Muestreo aleatorio.

Error muestral: Para el conjunto de la muestra, y a un nivel de confianza del 95%, el error es inferior al 1%.

Fecha de realización: Junio 2014.

3.2 VARIABLES

3.2.1 PROCEDIMIENTO DE SELECCIÓN DE LAS VARIABLES

En el presente trabajo se ha considerado un total de 34 variables o ratios financieros, obtenidos, todos ellos, de la revisión de la literatura previa sobre predicción de insolvencia.

Estas variables han sido seleccionadas entre aquellas variables con mayor frecuencia de utilización en los estudios previos de predicción de insolvencia (Bellovary et al., 2007). Por tanto, la gran mayoría de ellas son de tipo cuantitativo, correspondiéndose con distintos ratios económico-financieros obtenidos a partir de la información contable de las empresas pertenecientes a las distintas muestras.

Las variables se han clasificado en variables del modelo global (VMG) o descentrado y variables de los diferentes modelos de sectores económicos (VMS) o modelos centrados.

Adicionalmente, se han incorporado otras 5 variables explicativas de tipo cualitativo para recoger la pertenencia a alguno de los sectores económicos considerados en este estudio. Estas variables cualitativas han sido empleadas en la estimación de los modelos globales de predicción de insolvencia en cada una de las muestras.

Además de los ratios financieros objeto de análisis, se ha empleado una variable dummy, de tipo binomial, como variable dependiente al objeto de identificar a la empresa como solvente o como insolvente.

3.2.2 VARIABLES EMPLEADAS EN LOS MODELOS DESCENTRADOS

Las variables empleadas en los modelos descentrados (VMG) se obtuvieron seleccionando las diez variables con mayor frecuencia de utilización en los estudios previos (Bellovary et al., 2007). En concreto, la frecuencia de empleo de dichas variables en los estudios previos es la que aparece en la tabla 3.1.

TABLA 3.1 FRECUENCIA DE UTILIZACIÓN DE LOS RATIOS MÁS COMUNES EN LA LITERATURA PREVIA

Ratios	Frecuencia (estudios)
Beneficio Después de Impuestos/Total Activo	54
Activo Corriente/Pasivo Corriente	51
Fondo Maniobra/Total Activo	45
EBIT/Total Activo	42
Total Ingresos/Total Activo	35
Quick Ratio	32
Total Deuda/Total Activo	30
Activo Corriente/Total Activo	27
Beneficio Después de Impuestos/Patrimonio Neto	26

Fuente: Bellovary et al. (2007)

3.2.2 VARIABLES EMPLEADAS EN LOS MODELOS CENTRADOS

Para distinguir los diferentes sectores, la nomenclatura de las variables ha sido de VMSA para las variables específicas del sector agricultura, VMSI para las variables específicas del sector industrial, VMSC para las del sector de la construcción, VMSCS para las del sector comercio y servicios, y VMSH para el sector de hostelería.

Las variables específicas de cada uno de los sectores económicos considerados se seleccionaron a partir de la revisión de la literatura de aquellos estudios centrados en los sectores objeto de estudio. Del examen de las variables empleadas en los modelos, no se tuvo en consideración aquellas variables más frecuentemente empleadas en los modelos descentrados, seleccionando las variables que aparecieron como claramente específicas en la explicación de la insolvencia en los sectores económicos concretos y que hubieran aparecido, al menos, en más de un estudio.

En la tabla 3.2 puede observarse el detalle de todas las variables utilizadas en el presente trabajo, tanto las seleccionadas a partir de la revisión de la literatura de los modelos descentrados como las seleccionadas de los estudios orientados a sectores concretos o de modelos centrados.

TABLA 3.2 DESCRIPCIÓN DE LAS VARIABLES UTILIZADAS

Variables del Modelo Global (VMG)	
Beneficio Después de Impuestos/Total Activo	VMG1
Activo Corriente/Pasivo Corriente	VMG2
Fondo Maniobra/Total Activo	VMG3
EBIT/Total Activo	VMG4
Total Ingresos/Total Activo	VMG5
Quick Ratio	VMG6
Total Deuda/Total Activo	VMG7
Activo Corriente/Total Activo	VMG8
Beneficio Después de Impuestos/Patrimonio Neto	VMG9
Variables del Modelo del Sector Agricultura (VMSA)	
Fondos Propios/Total Deuda	VMSA1
EBIT/Gastos Financieros	VMSA2
EBIT/Total Ingresos	VMSA3
Variables del Modelo del Sector Industria (VMSI)	
Resultado de Explotación/Total Ingresos	VMSI1
Ventas/Cientes Medios	VMSI2
(Activo corriente-Pasivo corriente)/Total Capital	VMSI3
Fondos Propios/Pasivo no corriente	VMSI4
Gastos financieros/Total ingresos	VMSI5
LN (Activo)	VMSI6
Resultado de Explotación/Patrimonio Neto	VMSI7
Total Ingresos/Activo no corriente	VMSI8
Variables del Modelo del Sector Construcción (VMSC)	
Gastos financieros/EBIT	VMSC1
Resultado de Explotación/Total Ingresos	VMSC2
Fondos Propios/Total Deuda	VMSC3
Variables del Modelo de Comercio y Servicios (VMSCS)	
EBITDA/Total Pasivo	VMSCS1
EBIT/Gastos financieros	VMSCS2
EBIT/Pasivo corriente	VMSCS3
Ventas/Existencias	VMSCS4
Ventas/Total Activo	VMSCS5
Variables del Modelo de Hostelería (VMSH)	
EBITDA/Pasivo corriente	VMSH1
EBITDA/Total pasivo	VMSH2
Total Deuda Financiera/EBITDA	VMSH3
Total Deuda Financiera/Total Capital	VMSH4
Ventas a crédito/Cientes Medios	VMSH5
FCF/Total Deuda	VMSH6

En la tabla 3.3 se muestran tanto la descripción como la nomenclatura de las variables utilizadas en la investigación. Igualmente en la tabla 3.3 se detallan y definen las magnitudes contables que se incluyen dentro de cada partida utilizada en las variables seleccionadas, en base a la normativa del Plan General de Contabilidad actualmente en vigor en España, aprobado por los Reales Decretos 1514/2007 y 1515/2007.

TABLA 3.3 DEFINICIÓN DE LAS PARTIDAS INCLUIDAS EN LAS VARIABLES UTILIZADAS

Activo corriente	Activos vinculados al ciclo normal de explotación que la empresa espera vender, consumir o realizar en el transcurso del mismo (con carácter general el ciclo normal de explotación no excederá de un año) + aquellos activos no vinculados al proceso de explotación pero cuyo vencimiento, enajenación o realización se espera se realice en el corto plazo + efectivo y otros activos líquidos
Activo no corriente	Todos aquellos activos que no se consideren activos corrientes
Beneficio Después de Impuestos	Resultado neto de la empresa, una vez considerados todos los ingresos y todos los gastos.
Cash-flow libre	Representa el flujo de caja generado por la empresa, sin considerar la forma en que se ha financiado. Se calcularía como EBIT- Impuesto sobre EBIT + Depreciaciones y amortizaciones + Fondo de Maniobra - Inversiones en activo fijo
Cash-flow operativo	Beneficio neto + Amortizaciones
Deudas a largo plazo	Financiación ajena con vencimiento superior a un año
Deudores	Créditos con compradores de servicios que no tienen la condición estricta de clientes
EBIT	Beneficios antes de Intereses e Impuestos
EBITDA	Beneficios antes de Intereses, Impuestos, Amortizaciones y Depreciaciones

TABLA 3.3 DEFINICIÓN DE LAS PARTIDAS INCLUIDAS EN LAS VARIABLES UTILIZADAS (CONTINUACIÓN)

Fondo de Maniobra	Son los recursos de largo plazo que una empresa tiene para financiar las necesidades operativas de corto plazo, una vez que ha financiado sus activos fijos. Su cálculo sería (Deuda a largo plazo + Patrimonio Neto) - Activo Fijo
Gastos financieros	Coste de la financiación ajena
Ingresos	Ingresos de explotación
Inventario medio	$(\text{Existencias iniciales} + \text{Existencias Finales})/2$
Inversiones financieras a corto	Inversiones a corto plazo tanto de empresas del grupo, multigrupo, asociadas y otras partes vinculadas, como en aquellas empresas en las cuales no tenga la sociedad ninguna vinculación, incluyéndose adicionalmente los depósitos y fianzas
Margen de beneficio bruto	EBITDA/Ingresos de explotación
Pasivo corriente	Pasivos vinculados al ciclo normal de explotación que la empresa espera liquidar en el transcurso del mismo (con carácter general el ciclo normal de explotación no excederá de un año) + aquellos pasivos financieros clasificados como mantenidos para negociar, excepto los derivados financieros con vencimiento superior a un año

TABLA 3.3 DEFINICIÓN DE LAS PARTIDAS INCLUIDAS EN LAS VARIABLES UTILIZADAS (CONTINUACIÓN)

Pasivo no corriente	Toda aquella financiación ajena no considerada como pasivo corriente
Patrimonio Neto	Recursos propios de la empresa
Tesorería	Activos líquidos
Total Activo	Activo no corriente + Activo corriente.
Total Pasivo	Total financiación ajena de la empresa. Suma de Pasivo corriente y Pasivo no corriente.

BIBLIOGRAFÍA CAPITULO 3

Altman, D.G. (1991). *Practical Statistics for Medical Research*. Londres. Chapman.

Bolton, R. y D. Hand (2002). Statistical fraud detection: a review. *Statistical Science*, 17(3), 235-255.

Chye, K., Chin, T. y G. Peng (2004). Credit scoring using data mining techniques. *Singapore Management Review*, 26,2, 25-47

Fernández, M.A.; Gil, A.M. y P. López (2009). *Estructura y Equilibrio financiero de las PYMES Españolas ante las Nuevas Exigencias de Información Financiera*. Fundación de las Cajas de Ahorros. Madrid.

González, J., Perea, J. C. y A. Y. Sánchez (2003). El Reglamento sobre procedimientos de insolvencia de la Unión Europea. *Partida Doble*, 141, 6-15.

Hair, J. (2008). *Análisis Multivariante*. Editorial Prentice Hall. Madrid.

Ley 22/2003 de 9 de Julio, Concursal.

Ley 38/2011, de 10 de Octubre, de reforma de la Ley 22/2003, de 9 de Julio, Concursal.

López, M. y J. López (1996). *Estadística para actuarios*. Editorial Mapfre, Madrid.

Manual IBM SPSS Regression 2.0.

Olson, D. L., Delen, D. y Y. Meng (2012). Comparative analysis of data mining methods for bankruptcy prediction. *Decision Support Systems*, 52, 464- 473.

Real Decreto Ley 1514/2007 de 16 de Noviembre. Plan General de Contabilidad.

Real Decreto Ley 1515/2007 de 16 de Noviembre. Plan General de Contabilidad de pequeñas y medianas empresas.

Real Decreto Ley 3/2009 de 27 de Marzo, de medidas urgentes en materia tributaria.

Tsai, C.F. (2009). Feature selection in bankruptcy prediction. *Knowledge-Based Systems*, 22, 120–127.

Wang, C.-M. y Y.F. Huang (2009). Evolutionary-based feature selection approaches with new criteria for data mining: A case study of credit approval data. *Expert Systems with Applications*, 36, 5900–5908.

Zmijewski, M.E. (1984). Methodological Issues Related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22, 59-81.

CAPÍTULO 4:

RESULTADOS

4.1 INTRODUCCIÓN

Una vez detallados los datos y las variables a utilizar, se exponen en este capítulo los resultados del trabajo empírico realizado.

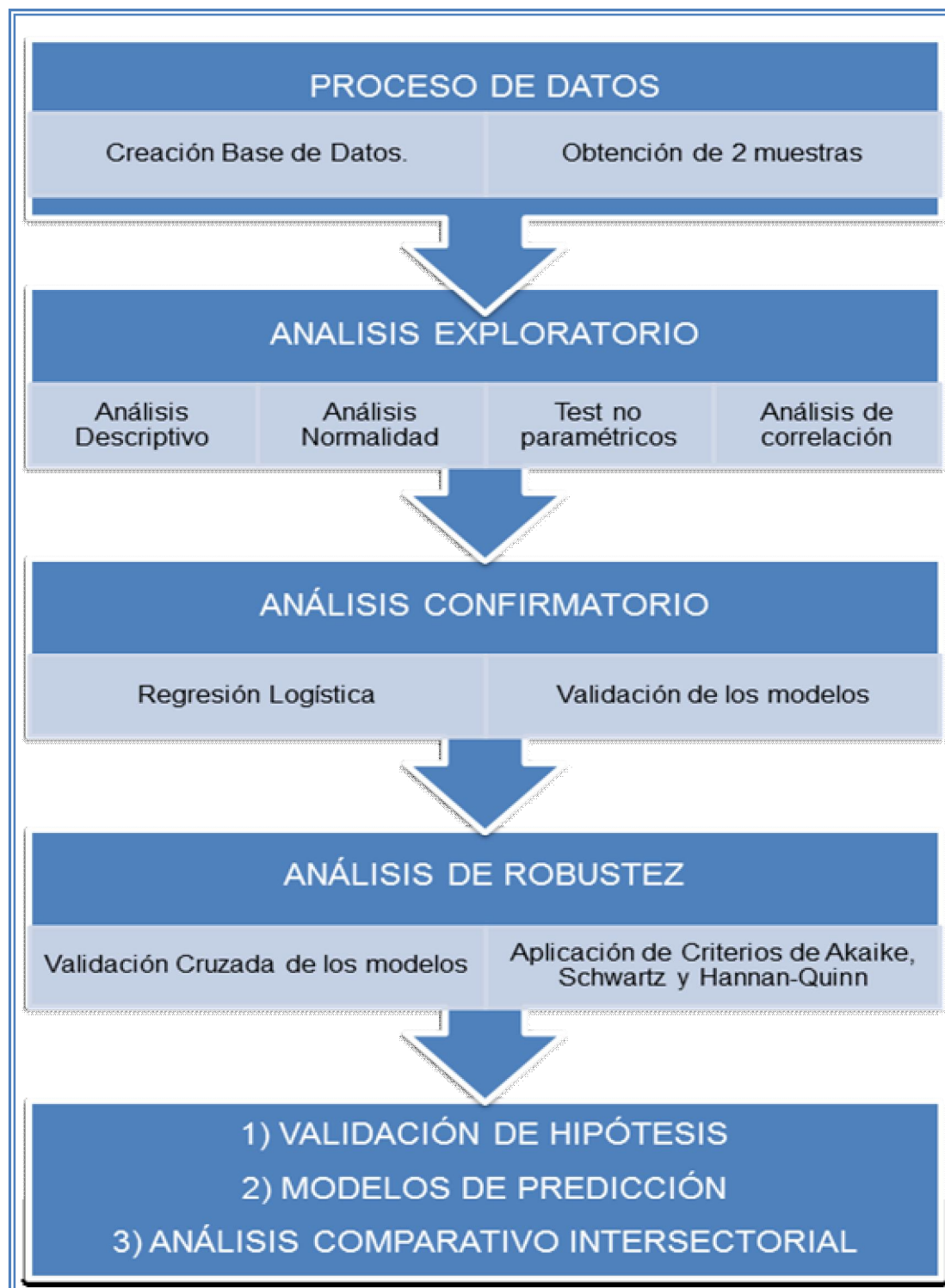
Para cada uno de los sectores económicos considerados en este trabajo –agricultura, industria, construcción, comercio y servicios, y hostelería- se ha llevado a cabo un análisis exploratorio de las variables seleccionadas al objeto de conocer sus principales parámetros estadísticos y comprobar que variables podrían resultar relevantes en la predicción de la insolvencia. Posteriormente, se ha realizado un análisis confirmatorio de los resultados revelados por el análisis descriptivo de las variables que permitirá identificar las variables predictoras de la solvencia en los diferentes sectores económicos analizados.

De la misma forma, se ha estimado un modelo global con una muestra compuesta con empresas de los cinco sectores económicos bajo estudio. Previamente, se ha llevado a cabo un análisis exploratorio de las variables seleccionadas.

Dado que en este trabajo se utilizan dos muestras con objeto de analizar la predicción de la insolvencia un año antes de la quiebra (muestra M.I) y dos años antes (muestra M.II), el análisis exploratorio y el confirmatorio se han efectuado, igualmente, para cada una de las muestras.

Un esquema de la metodología seguida se puede observar en el gráfico 4.1.

GRÁFICO 4.1 METODOLOGÍA APLICADA



4.2 RESULTADOS CON INFORMACIÓN DE UN AÑO ANTES DE LA QUIEBRA (M.I)

4.2.1 RESULTADOS DEL MODELO GLOBAL (M.I)

4.2.1.1 Análisis Exploratorio

En el análisis exploratorio se efectúa un estudio descriptivo con objeto de conocer los principales parámetros estadísticos de las variables utilizadas (media, mediana, desviación típica, mínimo y máximo). Posteriormente, se analizará la normalidad de las variables, lo que permitirá establecer si procede la aplicación de tests paramétricos o no paramétricos para, en último lugar, llevar a cabo un análisis de correlación entre las distintas variables.

A partir del análisis exploratorio, se podrá inducir, tentativamente, qué variables podrían resultar relevantes en la predicción de la solvencia en el modelo global. Este análisis será contrastado posteriormente mediante un análisis confirmatorio específico.

4.2.1.1.1 Análisis Descriptivo

El análisis descriptivo se presenta de forma diferenciada para las empresas solventes e insolventes de la muestra global de empresas. En esta muestra se encuentran incluidas empresas representativas de todos los sectores económicos considerados, un año antes de la entrada en quiebra. El objetivo de este análisis es la

comparación de los parámetros estadísticos obtenidos en cada variable, diferenciando según sean las empresas solventes o insolventes.

En la tabla 4.1 se detalla el análisis descriptivo realizado para la muestra global de empresas un año antes de la entrada en quiebra.

Observando las medias obtenidas para cada una de las variables, y diferenciando las empresas solventes de las empresas insolventes, se comprueba que en la muestra M.I, tanto VMG1 como VMG4 presentan medias con diferente signo según se trate de empresas solventes o insolventes.

Además, las medias y las medianas de cada una de las variables, en la mayor parte de los casos, presentan valores superiores en las empresas solventes. El 56% de los valores medios de las empresas solventes superan al de las empresas insolventes. Para el caso de la mediana este porcentaje es del 56%.

En concreto, las variables VMG3, VMG7 y VMG9 presentan medias inferiores en el caso de las empresas solventes frente a las insolventes.

De la observación de las desviaciones típicas, así como de los mínimos y máximos, se puede constatar la gran dispersión que existe en la mayor parte de los casos. Dispersión que, por otro lado, era previsible por las características propias de la información financiera, hecho

que ha sido constatado por diversos autores (Fernández et al., 2009).

Con respecto a la desviación típica, ésta presenta valores moderados, salvo para VMG2, VMG5 y VMG6. La desviación típica no debe analizarse de forma independiente, dado que para ser valorada correctamente han de tenerse en cuenta los valores habituales de la variable. En este sentido, siempre es adecuado considerarla conjuntamente con la media.

**TABLA 4.1 ESTADÍSTICOS DESCRIPTIVOS
(MODELO GLOBAL, M.I)**

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,05	-0,25	0,01	-0,04	0,21	1,00	-0,78	-19,99	5,26	0,53
	VMG2	3,37	1,81	1,30	0,91	11,68	4,67	0,00	0,00	242,72	64,52
	VMG3	0,24	0,32	0,21	0,30	0,30	0,44	-2,05	-2,49	0,97	1,00
	VMG4	0,05	-0,20	0,03	-0,02	0,41	0,72	-6,77	-13,36	5,28	0,59
	VMG5	2,14	1,33	1,22	0,87	11,34	1,89	0,00	0,00	324,76	36,09
	VMG6	2,34	0,74	0,85	0,48	10,25	1,22	0,00	0,00	218,42	14,13
	VMG7	0,23	0,42	0,13	0,33	0,28	0,47	0,00	0,00	2,33	5,70
	VMG8	0,60	0,59	0,63	0,65	0,29	0,31	0,00	0,00	1,00	1,00
	VMG9	0,04	0,30	0,05	0,07	1,95	2,39	-41,71	-14,95	7,91	37,41

En los Anexos de este trabajo se aportan los distintos histogramas de las variables donde se puede apreciar gráficamente la media y la dispersión de las mismas, tanto para empresas solventes como para las insolventes.

Aunque a partir de este análisis descriptivo se puede apreciar indicios de qué variables podrían resultar relevantes en el estudio de la insolvencia, aún no es posible llegar a concluir si estas variables son realmente significativas, por cuanto este primer análisis resultaría insuficiente para valorar si las diferencias presentadas se deben a su significación o a la variabilidad propia que presentan las variables económicas.

Este hecho implica la realización de distintos test que ayuden a valorar la significación de tales variables en el análisis de la insolvencia.

4.2.1.1.2 Análisis de Normalidad

A continuación conviene analizar la normalidad de las distintas variables. Este análisis es necesario para poder determinar cuáles serán los tests aplicables para el análisis de la significación de cada una de estas variables, de forma individual, en la predicción de la insolvencia. En el caso de resultar ser variables distribuidas normalmente, los tests a aplicar serán los tests paramétricos. Por el contrario, en caso de no distribuirse de acuerdo a una variable normal, los tests que se aplicarán serán los denominados no paramétricos.

De la misma forma, para el análisis de la normalidad también se diferenciará, en cada una de las muestras, los elementos correspondientes a las empresas solventes de las insolventes, debiendo cumplirse la normalidad en la

distribución en cada uno de los grupos para poder considerarla como una variable con distribución normal.

Para dilucidar esta cuestión, el test que se propone utilizar es el de Kolmogorov-Smirnov, cuya hipótesis nula es que los datos presentan una distribución normal. El cumplimiento del supuesto de normalidad implica que cada variable, en cada una de las muestras, y tanto para las empresas solventes como para las insolventes, se distribuye normalmente, con un nivel de significación igual o superior al 5%.

La tabla 4.2 muestran los resultados del test de Kolmogorov-Smirnov con un nivel de significación del 5%. Como se puede comprobar, todas las variables en M.I, tanto para empresas solventes como para insolventes, presentan un nivel de significación inferior al 5%. En consecuencia, se rechaza la hipótesis nula de normalidad.

El incumplimiento de la normalidad tiene dos implicaciones en el estudio empírico a realizar: Por un lado, para proseguir con el análisis descriptivo deberán aplicarse tests no paramétricos, como se ha apuntado anteriormente (utilizados para variables no distribuidas normalmente). En segundo lugar, y en tanto que se constata que no se cumple la normalidad, no sería recomendable realizar un análisis discriminante confirmatorio de ningún tipo, dado que es un requisito necesario para obtener resultados óptimos y robustos.

**TABLA 4.2 ANÁLISIS DE NORMALIDAD
(MODELO GLOBAL, M.I)**

		M.I (1 AÑO ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	0,325	0,000	0,330	0,000
	VMG2	0,383	0,000	0,397	0,000
	VMG3	0,071	0,000	0,088	0,000
	VMG4	0,346	0,000	0,279	0,000
	VMG5	0,425	0,000	0,235	0,000
	VMG6	0,406	0,000	0,432	0,000
	VMG7	0,211	0,000	0,184	0,000
	VMG8	0,093	0,000	0,113	0,000
	VMG9	0,393	0,000	0,225	0,000

NOTA: K-S: test de Kolmogorov-Smirnov

4.2.1.1.3 Tests no Paramétricos

A continuación, se aplicará a cada una de las variables distintos tests no paramétricos con objeto de constatar si estas variables son significativamente distintas entre las empresas solventes e insolventes. Hay que señalar que los tests no paramétricos son menos robustos que los tests paramétricos pero presentan, como principal ventaja, el no requerir la normalidad en los datos.

Se propone la aplicación de dos tests. Por un lado, se propone la aplicación del test de Mann-Whitney y, por otro, el test de Kolmogorov-Smirnov.

El contraste U de Mann-Whitney parte de la hipótesis de que dos poblaciones son independientes entre sí, donde la hipótesis nula a contrastar es que las dos poblaciones se

distribuyen de la misma manera. En caso de no aceptarse la hipótesis nula, esto implicaría un desplazamiento central de una de las distribuciones respecto a la otra, lo que sugeriría diferencia en la forma o en la dispersión de una población con respecto a la otra (López y López, 1996). Este test sería análogo al test "t" que se utiliza en los tests paramétricos.

Según Visauta (2003) el test de Kolmogorov-Smirnov para dos muestras sirve para contrastar si dos muestras proceden de poblaciones con la misma distribución, basándose en la comparación de sus funciones de distribución. La hipótesis nula, en este caso, es que no existen diferencias significativas entre las distribuciones de ambas muestras.

**TABLA 4.3 TESTS NO PARAMÉTRICOS
(MODELO GLOBAL, M.I)**

		M.I (1 AÑO ANTES)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	25,560	0,000	13,138	0,000
	VMG2	-12,047	0,000	5,687	0,000
	VMG3	5,130	0,000	3,862	0,000
	VMG4	-19,129	0,000	9,807	0,000
	VMG5	-5,872	0,000	2,691	0,000
	VMG6	-11,755	0,000	5,246	0,000
	VMG7	10,460	0,000	5,045	0,000
	VMG8	-0,112	0,911	1,449	0,030
	VMG9	2,381	0,017	5,325	0,000

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

Como puede observarse en la tabla 4.3, según el test de Mann Whitney, se rechaza la hipótesis nula en ocho de las nueve variables consideradas, aceptándose tan solo la igualdad de distribuciones en la variable VMG8. Esta variable rechazó la hipótesis nula con un nivel de significación del 5%.

Sin embargo, se rechaza la hipótesis nula de igualdad en las distribuciones mediante los resultados obtenidos por la aplicación del test de Kolmogorov-Smirnov en todas las variables consideradas en la muestra global.

Los resultados obtenidos en la muestra global para M.I en los distintos tests empleados son prácticamente idénticos, mostrando una clara significatividad en la práctica totalidad de las variables consideradas. Tan sólo la variable VMG8 no resultó ser significativa en el test de Mann-Whitney.

Por tanto, y dado el gran porcentaje de significación de casi la totalidad de las variables, parece posible la obtención de adecuados modelos de predicción, por tratarse de variables que a priori resultan ser apropiadas para valorar la insolvencia.

4.2.1.1.4 Análisis de Correlación

Para finalizar el análisis exploratorio, se procederá al cálculo de las correlaciones que se producen entre las variables, considerando cada pareja de variables posible,

incluida, propiamente, la variable insolvencia (que en nuestro caso será la futura variable dependiente).

En la tabla 4.4 se muestra el análisis de correlación bilateral para la muestra global correspondiente a M.I. Para analizar las correlaciones se ha utilizado el coeficiente de correlación de Pearson. Los valores de los distintos coeficientes estarán comprendidos entre -1 (una relación negativa perfecta) y $+1$ (una relación positiva perfecta). Un valor 0 indica que no existe una relación lineal entre las variables seleccionadas.

**TABLA 4.4 ANÁLISIS DE CORRELACIÓN BILATERAL
(MODELO GLOBAL, M.I.)**

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	INSOLV
VMG1	1	0,022	0,142	0,685	0,104	0,012	-0,103	0,051	-0,081*	-0,222
VMG2	0,022	1	0,085	0,030	-0,023	0,919	0,006	0,103	-0,009	-0,079
VMG3	0,142	0,085	1	0,133	-0,049	0,002	-0,014	0,502	-0,009	0,098
VMG4	0,685	0,030	0,133	1	0,119	0,016	-0,119	0,011	-0,086	-0,221
VMG5	0,104	-0,023	-0,049	0,119	1	-0,012	-0,064	0,066	-0,042	-0,046
VMG6	0,012	0,919	0,002	0,016	-0,012	1	-0,057	0,059	-0,005	-0,094
VMG7	-0,103	0,006	-0,014	-0,119	-0,064	-0,057*	1	-0,185	0,003	0,238
VMG8	0,051	0,103	0,502	0,011	0,066	0,059	-0,185	1	0,039	-0,011
VMG9	-0,081	-0,009	-0,009	-0,086	-0,042	-0,005	0,003	0,039	1	0,063
INSOLV	-0,222	-0,079	0,098	-0,221	-0,046	-0,094	0,238	-0,011	0,063	1

Este análisis permite no solo conocer a priori el signo de la relación entre variables sino el grado de relación existente entre las mismas. Se deberá, por tanto, tener especial cuidado en el análisis confirmatorio con aquellas

variables que presenten una elevada correlación, al objeto de no contar con variables redundantes en los modelos que puedan distorsionar los resultados obtenidos.

Prestando atención a los resultados de la tabla 4.4 puede comprobarse como existe una elevada correlación entre las variables VMG2 y VMG6, en concreto, con un coeficiente de correlación de Pearson de 0,919. La elevada correlación de esta pareja de variable era previsible ante la similitud en el cálculo de las mismas.

Como resumen del análisis exploratorio realizado, es posible concluir que de las nueve variables consideradas, prácticamente todas muestran signos de ser variables relevantes en el análisis de la solvencia.

4.2.1.2 Análisis Confirmatorio

Una vez concluido el análisis exploratorio para la muestra global en M.I., en este apartado se realizará un análisis confirmatorio de los resultados expuestos en el análisis descriptivo de las variables. En este sentido, se estimará un modelo global, sin incluir ninguna variable cualitativa representativa del sector de actividad.

En el capítulo 2 ya se detallaron los distintos modelos estadísticos aplicables a problemáticas como la que es objeto de este estudio, en la cual la variable dependiente o explicativa (la insolvencia) es una variable categórica de carácter dicotómico, que toma valores 0 ó 1.

Como ya se comentó anteriormente, el Análisis Discriminante no resulta aplicable, en tanto que no se verifica la normalidad de las variables seleccionadas, requisito indispensable y necesario para obtener un modelo de calidad.

Por otro lado, con respecto a la utilización de los modelos estadísticos Logit y Probit, y teniendo en cuenta que los resultados suelen ser similares en ambos modelos, hay que precisar que el modelo Logit cuenta con la ventaja añadida de favorecer una mejor interpretación de los resultados (aportando también una utilidad explicativa e interpretativa, más allá de la puramente predictiva). Por todo ello, se estimará un modelo Logit para cada una de las muestras, ya sean globales o sectoriales, tanto en M.I como en M.II.

Además, este tipo de modelo ha sido uno de los más asiduamente empleados en los estudios empíricos previos, dada su utilidad, alta capacidad predictiva y la posibilidad de interpretación de los resultados obtenidos.

Cada uno de los modelos estimados será validado en una muestra externa. Este procedimiento permitirá testar la capacidad de generalización de los modelos, elemento esencial en cualquier estudio de este tipo, en la medida en que su principal utilidad es poder utilizarlos en elementos no considerados para la estimación de los mismos o, lo que es lo mismo, para comprobar su capacidad de generalización.

La estructura de la muestra que se utilizará en cada caso para la estimación de los modelos será de un 70% de las empresas que formarán la denominada muestra de entrenamiento y el 30% restante para la denominada muestra de testeo.

A continuación, se estimará el modelo de regresión logística con la muestra global de empresas un año antes de su entrada en quiebra.

4.2.1.2.1 Modelo Logit Global sin variables cualitativas

Como ya se expuso en el capítulo 2, en el caso de los modelos Logit, se aplica el método de máxima verosimilitud de forma que se van realizando sucesivas iteraciones hasta obtener la estimación más probable.

En este sentido, existen varias formas de realizar estas iteraciones: hacia adelante (se van añadiendo variables y verificando el ajuste del modelo en cada caso), hacia atrás (se parte de un modelo con todas las variables y se van eliminando las menos significativas hasta el momento en que todas sean significativas, no pudiendo eliminarse ninguna) y en último lugar, se encuentra la opción Stepwise, que resultaría de una combinación de las dos técnicas anteriores. En este trabajo se ha aplicado la opción hacia atrás.

Se han realizado numerosas pruebas con las distintas variables que habían resultado ser las más

significativas en las iteraciones realizadas. Para ello, nos hemos ayudado del método de selección hacia atrás por pasos, en el cual el contraste para la eliminación se fundamenta en la probabilidad del estadístico de la razón de verosimilitud (estimaciones de máxima verosimilitud parcial).

La tabla 4.5 muestra los resultados de la estimación del modelo Logit global sin variables cualitativas obtenido para M.I. En este modelo han resultado ser cuatro las variables seleccionadas: VMG1, VMG3, VMG4 y VMG5.

Al analizar los signos de los coeficientes obtenidos para cada una de estas variables, se puede apreciar que las variables VMG1 y VMG5 presentan signo negativo, lo cual denota una relación inversa con la insolvencia. Por tanto, cuanto más alto son estos ratios menor es la probabilidad de que la empresa entre en quiebra.

En cambio, las variables VMG3 y VMG4 presentan signos positivos en sus coeficientes. En este caso, la relación con la insolvencia es directa, es decir, cuanto más alto es el ratio, mayor es la probabilidad de insolvencia.

Si bien los signos de los coeficientes son muy útiles para poder valorar la relación positiva o negativa con la variable dependiente, para evaluar e interpretar la relevancia de cada variable en el modelo estimado, resulta más esclarecedor analizar el Odds ratio de cada variable. Recordemos que el cálculo de estos Odds ratio son

necesarios para la adecuada interpretación de los resultados obtenidos en cada variable.

El Odds ratio indica el cambio relativo que experimenta el cociente de probabilidades ($p/1-p$) cuando la variable X_j aumenta una unidad (Levy y Varela, 2003). Si el coeficiente (Beta) es positivo, su transformación (antilogaritmo) será mayor que 1, y el Odds ratio aumentará. Este aumento se produce cuando la probabilidad prevista de ocurrencia de un suceso aumenta y la probabilidad prevista de su no ocurrencia disminuye. En consecuencia, el modelo tiene una elevada probabilidad de ocurrencia. De la misma forma, si el coeficiente es negativo, el antilogaritmo es menor que 1 y el Odds ratio disminuirá (Hair et al., 1999).

Observando los Odds ratio de las variables, como cabía esperar, tan solo son mayores que 1 los correspondientes a las variables VMG3 y VMG4, en tanto que son las únicas variables con signo positivo, presentando valores menores a 1 en el resto de los casos.

Así, en el caso concreto de VMG3, al tener signo positivo y un Odds ratio de 3,887 implica que cuando el ratio "Fondo de Maniobra/Total Activo" aumenta en un punto, la probabilidad de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 3,887 veces. Las variables con signo negativo (VMG1 y VMG5) presentan un Odds ratio entre 0 y 1. En este caso, el Odds ratio para VMG1 tiene un valor de 0,000 lo cual indica que, si bien con respecto a la insolvencia no es una

variable relevante, sí lo es con respecto a las empresas solventes.

En cuanto a la significatividad de las variables incluidas en el modelo, todas las variables presentan niveles de significatividad al 5%, siendo, por tanto, relevantes en la explicación del modelo.

Una vez valorada la significación de cada variable de forma individual, es el momento de analizar distintos tests para constatar la robustez y precisión del modelo en su conjunto.

En primer lugar, la prueba de Razón de Verosimilitud (RV) de todos los coeficientes, o también denominado prueba Ómnibus, que estudia la significatividad conjunta de las variables seleccionadas. En el caso del modelo global en M.I el valor obtenido es de 0,000, concluyéndose que el ajuste es adecuado, dado que al menos uno de los predictores se encuentra significativamente relacionado con la variable dependiente.

En segundo lugar, la matriz de clasificación, que evalúa la capacidad predictiva del modelo, muestra que se han clasificado correctamente el 80,1% de las empresas, obteniéndose un mayor nivel de predicción en el caso de las empresas solventes, con un 86,2% de acierto, y clasificando correctamente el 73,7% de las empresas insolventes.

Por su parte, el nivel de clasificación en la muestra externa, utilizada para comprobar la capacidad de generalización del modelo, ha resultado ser del 80,8%.

TABLA 4.5 MODELO LOGIT GLOBAL SIN VARIABLES CUALITATIVAS. M.I

Ajuste del modelo		Valor	
Significatividad Prueba Ómnibus		0,000	
Significatividad Test de Hosmer y Lemeshow		0,085	
-2 log verosimilitud (-2LL)		390,738	
R ² Cox & Snell		0,331	
R ² Nagelkerke		0,441	
Matriz de clasificación		%	
Dentro de la muestra		80,1	
	Insolventes	73,7	
	Solventes	86,2	
Muestra de testeo		80,8	
	Insolventes	74,5	
	Solventes	86,9	
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-21,653	0,000	0,000
VMG3	1,358	3,887	0,001
VMG4	7,758	2.339,384	0,033
VMG5	-0,286	0,751	0,002
Constante	-0,460	0,631	0,043

Notas: VMG1= Beneficio después de Impuestos/Total Activo; VMG3=Fondo de Maniobra/Total Activo; VMG4= EBIT/Total Activo; VMG5= Total Ingresos/Total Activo.

En general, ante un acontecimiento de respuesta binaria en los que la respuesta puede ser positiva (p) o negativa (n), (como es en el presente caso en el que las

empresas pueden ser solventes o insolventes), se pueden dar cuatro resultados distintos según si el valor predicho coincide con el valor real. Así, si una empresa es solvente y se predice que la empresa es solvente, se considerará como un verdadero positivo (VP). En cambio, si siendo la empresa solvente, se predice como insolvente, se habrá producido lo que se denomina un falso positivo (FP).

Por análogo razonamiento, si la empresa es insolvente y se predice como solvente, se obtendrá lo que se conoce como un falso negativo (FN). Igualmente, se tendrá un verdadero negativo (VN) en caso de ser la empresa insolvente y se predice como tal. En la tabla 4.6 se muestran estas cuatro posibilidades.

TABLA 4.6 RESULTADOS POSIBLES EN UNA MATRIZ DE CLASIFICACIÓN

		VALOR REAL	
		Solvente	Insolvente
PREDICCIÓN	Solvente	Verdaderos Positivos (VP)	Falsos Positivos (FP)
	Insolvente	Falsos Negativos (FN)	Verdaderos Negativos (VN)

Con objeto de evaluar la bondad del ajuste del modelo se utiliza, además, el test denominado de Hosmer y Lemeshow. Este estadístico se fundamenta en agrupar los casos en deciles de riesgo y comparar la probabilidad observada con la probabilidad esperada dentro de cada decil. Si el nivel de significatividad es superior a 0,05 el modelo contará con un buen ajuste, en la medida en que se acepta la hipótesis nula de que no existen diferencias

entre los valores observados y los valores predichos por el modelo. En nuestro caso, para M.I, la significatividad del estadístico es 0,085 y, en consecuencia, se aceptaría la hipótesis de que no existen diferencias significativas entre las clasificaciones observadas y predichas.

Por otro lado, el test $-2 \log$ de verosimilitud ($-2LL$) es una medida global del ajuste del modelo. Un modelo con buen ajuste tendrá un valor pequeño para $-2LL$ (Hair et al., 1999). Para M.I, toma un valor de 390,738.

El coeficiente de Cox y Snell estima la proporción de la varianza de la variable dependiente explicada por las variables independientes o predictoras. Su valor se encuentra comprendido entre 0 y valores inferiores a 1, si bien el nivel máximo, 1, nunca llega a alcanzarse. Para M.I, asciende a 0,331.

Por último, el R^2 de Nagelkerke es una modificación del coeficiente de Cox y Snell, corrigiendo la escala del estadístico para cubrir el rango completo de 0 a 1. De esta forma, además de que su valor suele ser superior, es un estadístico más fácil de interpretar. Para este caso alcanza un valor de 0,441. Ambos resultados denotan un ajuste aceptable del modelo seleccionado.

4.2.2 RESULTADOS MODELO AGRICULTURA

4.2.2.1 Análisis Exploratorio

Al igual que para la muestra global de empresas pertenecientes a todos los sectores económicos, en este apartado se aborda el análisis exploratorio para la muestra de empresas pertenecientes al sector agricultura.

Se analizarán los principales parámetros estadísticos de las variables utilizadas (media, mediana, desviación típica, mínimo y máximo) para, a continuación, analizar la normalidad de las mismas. Esto permitirá determinar si resulta conveniente la aplicación de tests paramétricos o no paramétricos y, posteriormente, se llevará a cabo un análisis de correlación entre las distintas variables.

4.2.2.1.1 Análisis Descriptivo

Este análisis se lleva a cabo con objeto de comparar los parámetros estadísticos obtenidos en cada variable, según las empresas sean solventes o insolventes.

En la tabla 4.7 se detalla el análisis descriptivo realizado para la muestra del sector de agricultura un año antes de la quiebra, M.I. Observando las medias obtenidas para cada una de las variables, y diferenciando las empresas solventes de las empresas insolventes, se comprueba que sólo VMG1, VMG4 y VMSA2 presentan medias con diferente signo según se trate de empresas solventes e insolventes.

Considerando los valores de las medias y las medianas, puede apreciarse como en M.I, el 67% de los valores medios en las empresas solventes supera al de las empresas insolventes. Para el caso de la mediana, este porcentaje alcanza el 75%.

**TABLA 4.7 ESTADÍSTICOS DESCRIPTIVOS
(MODELO AGRICULTURA, M.I)**

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,02	-0,03	0,01	-0,01	0,03	0,07	-0,03	-0,19	0,15	0,13
	VMG2	1,55	0,99	1,22	0,84	0,97	0,87	0,02	0,03	4,25	4,47
	VMG3	0,20	0,27	0,14	0,16	0,28	0,28	-0,62	-0,22	0,69	0,80
	VMG4	0,04	-0,01	0,03	0,00	0,04	0,08	-0,00	-0,19	0,20	0,19
	VMG5	0,93	0,66	0,76	0,31	0,67	0,84	0,09	0,01	2,77	3,52
	VMG6	1,03	0,46	0,67	0,34	0,95	0,42	0,02	0,01	3,13	1,61
	VMG7	0,33	0,47	0,34	0,54	0,21	0,30	0,00	0,00	0,73	1,24
	VMG8	0,45	0,46	0,36	0,49	0,28	0,29	0,01	0,01	0,98	0,97
	VMG9	0,08	0,16	0,04	0,00	0,07	0,57	0,01	-1,03	0,26	1,55
VARIABLES MODELO SECTORIAL (AGRICULTURA)	VMSA1	2,87	0,75	1,05	0,11	4,56	2,20	-0,32	-2,04	22,98	8,50
	VMSA2	4,99	-1,14	1,76	0,28	13,18	17,64	-0,28	-63,57	76,35	64,38
	VMSA3	0,08	-0,02	0,04	0,01	0,17	0,29	-0,01	-0,77	1,00	0,68

En concreto, las variables VMG3, VMG7 y VMG9 son las variables que presentan medias inferiores en el caso de que las empresas sean solventes o insolventes.

Con respecto a la desviación típica, esta presenta valores moderados, lo que indica un grado de dispersión satisfactorio.

4.2.2.1.2 Análisis de Normalidad

A continuación, procede analizar la normalidad de las distintas variables.

En la tabla 4.8 se muestran los resultados del test de Kolmogorov-Smirnov para las empresas solventes e insolventes pertenecientes al sector agricultura, con un nivel de significación del 5%.

**TABLA 4.8 ANÁLISIS DE NORMALIDAD
(SECTOR AGRICULTURA, M.I)**

		M.I (1 AÑO ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	0,194	0,000	0,300	0,000
	VMG2	0,280	0,000	0,255	0,000
	VMG3	0,134	0,006	0,082	0,200
	VMG4	0,161	0,000	0,268	0,000
	VMG5	0,159	0,000	0,184	0,000
	VMG6	0,256	0,000	0,303	0,000
	VMG7	0,168	0,000	0,263	0,000
	VMG8	0,124	0,016	0,149	0,001
	VMG9	0,193	0,000	0,173	0,000
VARIABLES MODELO SECTORIAL (AGRICULTURA)	VMSA1	0,246	0,000	0,366	0,000
	VMSA2	0,504	0,000	0,506	0,000
	VMSA3	0,417	0,000	0,447	0,000

NOTA: K-S: test de Kolmogorov-Smirnov

Con la excepción de VMG3, para el caso de las empresas insolventes, todas las variables, tanto para empresas solventes como insolventes, presentan

significación al nivel del 5%. En consecuencia, se rechazaría la hipótesis nula de normalidad. El incumplimiento de la normalidad implica necesariamente la aplicación de tests no paramétricos.

4.2.2.1.3 Tests no paramétricos

Según los resultados que aparecen en la tabla 4.9, se acepta la hipótesis nula de igualdad de distribuciones para VMG5 y VMG8, que presentan tanto el test de Mann-Whitney como el de Kolmogorov-Smirnov, con niveles de significación superiores al 5%.

**TABLA 4.9 TESTS NO PARAMÉTRICOS
(SECTOR AGRICULTURA, M.I)**

		M.I (1 AÑO ANTES)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	-6,565	0,000	3,794	0,000
	VMG2	-4,724	0,000	2,455	0,000
	VMG3	1,765	0,078	1,108	0,171
	VMG4	-5,329	0,000	3,013	0,000
	VMG5	-0,203	0,840	0,650	0,792
	VMG6	-3,832	0,000	2,031	0,001
	VMG7	3,298	0,001	1,556	0,016
	VMG8	0,039	0,969	0,583	0,886
	VMG9	-0,297	0,766	2,505	0,000
VARIABLES MODELO SECTORIAL (AGRICULTURA)	VMSA1	-5,274	0,000	2,905	0,000
	VMSA2	-6,157	0,000	3,368	0,000
	VMSA3	-4,913	0,000	2,784	0,000

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

Por su parte, VMG3 y VMG9 sólo muestran niveles de significatividad superiores al 5% en el test de Mann-Whitney.

Con las excepciones apuntadas, se rechaza la hipótesis nula de igualdad en las distribuciones de las variables consideradas en el sector agricultura.

A partir de los resultados obtenidos, y dada la significación de casi la totalidad de las variables, parece posible la obtención de adecuados modelos de predicción, por tratarse de variables que a priori resultan ser apropiadas para valorar la insolvencia.

4.2.2.1.4 Análisis de Correlación

A continuación, se procederá al cálculo de las correlaciones que se producen entre las variables, considerando cada pareja de variables posible, incluida, propiamente, la variable insolvencia.

Respecto a los resultados del análisis de correlación bilateral para la muestra correspondiente al sector agricultura un año antes de la quiebra (tabla 4.10), se detecta una elevada correlación entre las variables VMG1 y VMG4, que presentan, en concreto, un coeficiente de correlación de Pearson de 0,993. Asimismo, la correlación es elevada entre las parejas VMG2 y VMG6 (0,905) y entre VMG1 y VMG7 (-0,867).

TABLA 4.10 ANÁLISIS DE CORRELACIÓN BILATERAL
(SECTOR AGRICULTURA, M.I)

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	VMSA1	VMSA2	VMSA3	INSOLV
VMG1	1	0,139	0,069	0,99	-0,172	0,121	-0,867	0,116	-0,215	0,161	0,430	0,285	-0,302
VMG2	0,139	1	0,007	0,120	-0,193	0,905	-0,129	0,114	-0,075	0,442	0,199	0,110	-0,319
VMG3	0,069	0,007	1	0,067	0,210	-0,088	-0,042	0,540	0,130	-0,084	-0,028	-0,059	0,142
VMG4	0,993	0,120	0,067	1	-0,074	0,115	-0,862	0,132	-0,172	0,158	0,512	-0,059	-0,284
VMG5	-0,172	-0,193	0,210	-0,074	1	-0,175	0,122	0,427	0,252	-0,118	-0,024	0,304	0,052
VMG6	0,121	0,905	-0,088	0,115	-0,175	1	-0,111	0,066	-0,055	0,561	0,317	-0,091	-0,289
VMG7	-0,867	-0,129	-0,042	-0,862	0,122	-0,111	1	-0,156	0,212	-0,109	-0,060	-0,110	0,247
VMG8	0,116	0,114	0,540	0,132	0,427	0,066	-0,156	1	0,154	-0,147	0,185	0,104	0,003
VMG9	-0,215	-0,075	0,130	-0,172	0,252	-0,055	0,212	0,154	1	-0,245	0,113	0,094	0,160
VMSA1	0,161	0,442	-0,084	0,158	-0,118	0,561	-0,109	-0,147	-0,245	1	0,031	0,109	-0,103
VMSA2	0,430	0,199	-0,028	0,512	-0,024	0,317	-0,060	0,185	0,113	0,031	1	0,885	0,096
VMSA3	0,285	0,110	-0,059	0,304	-0,091	0,156	-0,110	0,104	0,094	0,109	0,885	1	0,034
INSOLV	-0,302	-0,319	0,142	-0,284	0,052	-0,289	0,247	0,003	0,160	-0,103	0,096	0,034	1

Como resumen del análisis exploratorio realizado, es posible concluir que, de las doce variables consideradas, prácticamente todas muestran signos de ser variables relevantes en el análisis de la solvencia.

4.2.2.2 Análisis Confirmatorio

En este apartado se realizará un análisis confirmatorio de dichos resultados. En este sentido, se estimará un modelo Logit para la muestra seleccionada del sector agricultura al objeto de predecir la insolvencia con información de un año antes de producirse la quiebra.

El modelo Logit estimado permitirá interpretar los resultados y obtener conclusiones de cuáles son las variables más relevantes en la predicción de la insolvencia en el sector agricultura. Al igual que con el modelo global, este modelo se testará en una muestra externa. Esto permitirá validar la capacidad de generalización del mismo. Para tal fin, la muestra se dividirá en dos partes: un 70% para la denominada muestra de entrenamiento y un 30% para la denominada muestra de testeo.

4.2.2.2.1 Modelo Logit Sector Agricultura

Como ya se expuso anteriormente, se aplica el método de máxima verosimilitud, de forma que se irán realizando sucesivas iteraciones hasta obtener la estimación más probable.

En la tabla 4.11 se exponen los resultados del modelo Logit estimado para la muestra de empresas del sector agricultura, un año antes de la quiebra, M.I., y que ha resultado ser la regresión logística que mejores resultados presentaba en lo que respecta a nivel de predicción. Las variables seleccionadas han sido tres en el modelo elegido: VMG1, VMG6 y VMSA2.

Al analizar los signos de los coeficientes obtenidos para cada una de estas variables, se puede apreciar que las variables VMG1 y VMG6 presentan signo negativo, lo cual denota una relación inversa con la insolvencia. En consecuencia, cuanto más alto sean estos ratio menor es la probabilidad de que la empresa entre en quiebra.

En cambio, la variable VMSA2 presenta signo positivo en su coeficiente. En este caso, la relación con la insolvencia es directa, es decir, cuanto más alto el ratio mayor es la probabilidad de insolvencia.

Para valorar e interpretar la relevancia de cada variable en el modelo estimado resulta más esclarecedor analizar el Odds ratio de cada variable. Como cabía esperar, tan solo es ligeramente superior a 1 el correspondiente a la variable VMSA2, en tanto que para VMG1 resulta inferior a 1, en consonancia con el signo que presenta el coeficiente de esta variable.

En el caso de VMSA2, al tener signo positivo y un Odds Ratio de 1,004 expresa que cuando el ratio "Beneficio antes de intereses e impuestos/Gastos

financieros" aumenta en un punto, la probabilidad de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 1,004 veces.

Como puede comprobarse, las tres variables son significativas al 5%.

**TABLA 4.11 MODELO LOGIT
SECTOR AGRICULTURA. M.I**

Ajuste del modelo		Valor	
Significatividad prueba Ómnibus		0,000	
Significatividad Test de Hosmer y Lemeshow		0,487	
-2 log verosimilitud (-2LL)		78,310	
R ² Cox & Snell		0,420	
R ² Nagelkerke		0,560	
Matriz de clasificación		%	
Dentro de la muestra		78,5	
	Insolventes	78,3	
	Solventes	78,7	
Muestra de testeo		75,6	
	Insolventes	90,9	
	Solventes	57,9	
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-32,529	0,000	0,002
VMG6	-0,973	0,378	0,032
VMSA2	0,004	1,004	0,003
Constante	0,753	2,123	0,073

NOTA: VMG1 = Beneficio después de Impuestos/Total Activo; VMG6 = Quick Ratio; VMSA2 = EBIT/Gastos Financieros.

Una vez valorada la significación de cada variable de forma individual, procede analizar distintos tests para

constatar la robustez y precisión del modelo en su conjunto.

La prueba Ómnibus ofrece un valor de 0,000, lo que indicaría un ajuste adecuado, dado que al menos uno de los predictores está significativamente relacionado con la variable dependiente.

En segundo lugar, mediante la matriz de clasificación, la cual evalúa la capacidad predictiva del modelo, se observa que se han clasificado correctamente el 78,5% de las empresas, obteniéndose un mayor nivel de predicción en el caso de las empresas solventes, con un 78,7% de acierto, y clasificando correctamente el 78,3% de las empresas insolventes.

Por su parte, el nivel de clasificación en la muestra externa, utilizada para comprobar la capacidad de generalización del modelo, ha resultado ser del 75,6%.

La bondad del ajuste mediante el test de Hosmer y Lemeshow arroja una significatividad del estadístico de 0,487 y, en consecuencia, se aceptaría la hipótesis de que no existen diferencias significativas entre las clasificaciones observadas y predichas.

Por otro lado el test $-2 \log$ de verosimilitud ($-2LL$) toma un valor de 78,310.

Finalmente, el coeficiente de Cox y Snell y el R^2 de Nagelkerke alcanzan valores de 0,420 y 0,560,

respectivamente. Ambos resultados denotan también un ajuste aceptable del modelo seleccionado.

4.2.3 RESULTADOS MODELO INDUSTRIA

4.2.3.1 Análisis Exploratorio

En este apartado se abordará el análisis exploratorio para la muestra de empresas pertenecientes al sector industria utilizando información de un año antes de la entrada en quiebra.

4.2.3.1.1 Análisis Descriptivo

La tabla 4.12 ofrece un detalle del análisis descriptivo realizado para la muestra del sector industrial un año antes de la entrada en quiebra. De las medias obtenidas para cada una de las variables, y diferenciando las empresas solventes de las empresas insolventes, se comprueba que VMG1, VMG4, VMSI1 y VMSI3 presentan medias con diferente signo, según se trate de empresas solventes e insolventes.

Si se analizan los valores de las medias, se puede apreciar cómo, en M.I, el 41% de los valores medios en las empresas solventes supera al de las empresas insolventes. Para el caso de la mediana este porcentaje es del 47%.

En concreto, las variables VMG3, VMG7, VMG8, VMSI3, VMSI5 y VMSI7 presentan medias inferiores en el caso de las empresas solventes frente a las insolventes.

**TABLA 4.12 ESTADÍSTICOS DESCRIPTIVOS
(MODELO INDUSTRIA, M.I)**

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,03	-0,10	0,01	0,02	0,05	0,21	-0,07	-0,81	0,28	0,18
	VMG2	1,75	1,26	1,34	1,05	1,40	1,10	0,18	0,07	8,93	5,92
	VMG3	0,26	0,47	0,25	0,49	0,21	0,27	-0,31	-0,18	0,83	0,91
	VMG4	0,05	-0,06	0,03	0,00	0,06	0,19	-0,09	-0,73	0,39	0,34
	VMG5	1,37	1,32	1,20	1,15	0,75	0,93	0,18	0,00	4,05	4,39
	VMG6	1,14	0,82	0,81	0,68	1,20	0,80	0,05	0,01	8,61	4,86
	VMG7	0,27	0,30	0,25	0,19	0,19	0,31	0,00	0,00	0,87	1,54
	VMG8	0,53	0,57	0,54	0,58	0,22	0,23	0,08	0,02	0,94	0,96
	VMG9	0,10	0,08	0,05	0,48	0,15	1,09	-0,27	-5,04	0,80	3,08
VARIABLES MODELO SECTORIAL (INDUSTRIA)	VMSI1	0,04	-0,12	0,03	0,00	0,05	0,37	-0,05	-1,67	0,32	0,66
	VMSI2	14,28	8,33	5,75	4,11	29,18	14,47	1,47	0,01	260,75	91,93
	VMSI3	-0,09	0,07	0,38	0,54	2,09	3,06	-14,73	-17,59	4,12	5,16
	VMSI4	27,49	7,49	2,30	0,44	82,89	39,01	-0,96	-29,39	593,23	247,31
	VMSI5	0,01	0,02	0,00	0,02	0,02	0,02	0,00	0,00	0,11	0,09
	VMSI6	7,36	6,71	7,23	6,64	1,81	0,98	3,16	4,27	12,76	8,80
	VMSI7	0,20	0,35	0,12	0,18	0,31	0,96	-0,84	-3,02	2,21	3,72
	VMSI8	4,75	4,75	2,87	3,05	5,15	4,80	0,22	0,04	29,86	21,28

De la observación de las desviaciones típicas, así como de los mínimos y máximos, se puede constatar la gran dispersión que existe en la mayor parte de los casos. Con respecto a la desviación típica, presenta valores moderados, salvo en el caso de VMSI2 y VMSI4.

4.2.3.1.2 Análisis de Normalidad

Para el análisis de la normalidad, también se diferenciará en cada una de las muestras los elementos correspondientes a empresas solventes de los insolventes, debiendo cumplirse la normalidad en la distribución en cada uno de los grupos para así poder considerarla como una variable con distribución normal.

**TABLA 4.13 ANÁLISIS DE NORMALIDAD
(MODELO INDUSTRIA, M.I)**

		M.I (1 AÑO ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	0,279	0,000	0,342	0,000
	VMG2	0,325	0,000	0,311	0,000
	VMG3	0,050	0,200	0,161	0,000
	VMG4	0,231	0,000	0,265	0,000
	VMG5	0,188	0,000	0,149	0,000
	VMG6	0,296	0,000	0,268	0,000
	VMG7	0,155	0,000	0,272	0,000
	VMG8	0,066	0,026	0,132	0,000
	VMG9	0,453	0,000	0,317	0,000
VARIABLES MODELO SECTORIAL (INDUSTRIA)	VMSI1	0,186	0,000	0,362	0,000
	VMSI2	0,369	0,000	0,499	0,000
	VMSI3	0,268	0,000	0,392	0,000
	VMSI4	0,362	0,000	0,406	0,000
	VMSI5	0,250	0,000	0,408	0,000
	VMSI6	0,072	0,010	0,052	0,200
	VMSI7	0,425	0,000	0,392	0,000
	VMSI8	0,340	0,000	0,435	0,000

Los resultados del test de Kolmogorov-Smirnov para las empresas solventes e insolventes pertenecientes al

sector industria, con un nivel de significación del 5%, aparecen en la tabla 4.13. Con la excepción de VMG3, para el caso de las empresas solventes, y de VMSI6, para las empresas insolventes, todas las variables en M.I son significativas al nivel del 5%. Esto implica rechazar la hipótesis nula de normalidad de los datos y, por tanto, supone la aplicación de tests no paramétricos para discernir si existen diferencias entre los datos de las empresas solventes e insolventes.

4.2.3.1.3 Tests no paramétricos

En la tabla 4.14 aparecen los resultados de la aplicación de los mencionados tests a la muestra correspondiente a un año antes de la quiebra. Se observa que tan solo VMG5 registra un test de Mann-Whitney con un nivel de significación superior al 5% ($p > 0,05$), con lo cual se aceptaría, sólo en este caso, la hipótesis nula de igualdad de distribuciones para esta variable.

Con la excepción apuntada, se rechaza la hipótesis nula de igualdad en las distribuciones mediante los resultados obtenidos por la aplicación del test de Mann-Whitney y de Kolmogorov-Smirnov en todas las variables consideradas en el sector industria.

**TABLA 4.14 TESTS NO PARAMÉTRICOS
(MODELO INDUSTRIA, M.I)**

		M.I (1 AÑO ANTES)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	-10,994	0,000	6,049	0,000
	VMG2	-7,442	0,000	3,304	0,000
	VMG3	3,911	0,000	2,822	0,000
	VMG4	-9,353	0,000	5,400	0,000
	VMG5	0,174	0,862	1,887	0,002
	VMG6	-4,491	0,000	1,940	0,001
	VMG7	-3,164	0,002	2,214	0,000
	VMG8	2,651	0,008	2,001	0,001
	VMG9	2,302	0,021	3,326	0,000
VARIABLES MODELO SECTORIAL (INDUSTRIA)	VMSI1	-9,766	0,000	5,513	0,000
	VMSI2	-4,578	0,000	2,733	0,000
	VMSI3	2,955	0,003	2,493	0,000
	VMSI4	-7,263	0,000	3,235	0,000
	VMSI5	3,591	0,000	2,217	0,000
	VMSI6	-3,692	0,000	2,098	0,000
	VMSI7	2,433	0,015	3,089	0,000
	VMSI8	1,257	0,000	1,755	0,004

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

4.2.3.1.4 Análisis de Correlación

Para finalizar el análisis exploratorio del sector industrial en M.I, se procederá al cálculo de las correlaciones que se producen entre las variables.

El análisis de correlación bilateral para la muestra correspondiente a M.I. del sector industrial aparece en la tabla 4.15. Para analizar las correlaciones se ha empleado el coeficiente de correlación de Pearson. Puede

comprobarse cómo existe una elevada correlación entre los pares de variables VMG2 y VMG6 (0,902), VMG9 y VMSI7 (0,795) y VMSI3 y VMSI7 (0,794).

**TABLA 4.15 ANÁLISIS DE CORRELACIÓN BILATERAL
(MODELO INDUSTRIA, M.I.)**

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	VMSI1	VMSI2	VMSI3	VMSI4	VMSI5	VMSI6	VMSI7	VMSI8	INSOLV
VMG1	1	0,0803	0,208	0,541	-0,186	0,085	-0,165	0,099	-0,071	0,061	-0,484	-0,007	0,065	-0,028	0,146	-0,012	-0,006	-0,204
VMG2	0,083	1	0,107	0,118	-0,123	0,90	-0,105	0,173	-0,021	0,062	-0,029	0,008	0,289	-0,052	0,021	-0,020	-0,018	-0,181
VMG3	0,208	0,107	1	0,338	-0,103	0,077	-0,143	0,467	0,019	0,188	-0,184	0,028	0,030	0,067	0,102	0,051	0,053	0,083
VMG4	0,541	0,118	0,338	1	-0,297	0,118	-0,162	0,009	-0,119	0,123	-0,475	-0,019	0,137	-0,045	0,306	-0,033	-0,022	-0,310
VMG5	-0,186	-0,123	-0,103	-0,297	1	-0,110	-0,106	0,124	0,010	0,080	0,093	0,055	-0,030	-0,114	-0,395	0,043	0,066	0,069
VMG6	0,085	0,902	0,077	0,118	-0,110	1	-0,092	0,184	-0,031	0,066	-0,031	0,046	0,154	-0,049	0,037	-0,020	-0,008	-0,180
VMG7	-0,165	-0,105	-0,143	-0,162	-0,106	-0,092	1	-0,272	0,004	-0,060	-0,040	-0,036	-0,238	0,033	0,082	-0,021	-0,068	-0,002
VMG8	0,099	0,173	0,467	0,009	0,124	0,184	-0,272	1	0,050	0,092	-0,116	0,144	0,099	0,066	-0,128	0,060	0,172	0,096
VMG9	-0,071	-0,021	0,019	-0,119	0,010	-0,031	0,004	0,050	1	-0,008	0,005	0,660	-0,036	-0,007	-0,024	0,795	0,030	0,107
VMSI1	0,061	0,062	0,188	0,123	0,080	0,066	-0,060	0,092	-0,008	1	-0,031	0,000	0,041	-0,169	0,050	0,004	0,013	-0,151
VMSI2	-0,484	-0,029	-0,184	-0,475	0,093	-0,031	-0,040	-0,116	0,005	-0,031	1	0,005	0,025	-0,013	-0,111	-0,001	-0,007	0,049
VMSI3	-0,007	0,008	0,028	-0,019	0,055	0,046	-0,036	0,144	0,660	0,000	0,005	1	0,021	0,000	-0,018	0,794	0,009	0,059
VMSI4	0,065	0,289	0,030	0,137	-0,030	0,154	-0,238	0,099	-0,036	0,041	0,025	0,021	1	-0,066	0,070	-0,065	0,022	-0,166
VMSI5	-0,028	-0,052	0,067	-0,045	-0,114	-0,049	0,033	0,066	-0,007	-0,169	-0,013	0,000	-0,066	1	-0,032	-0,004	-0,018	0,107
VMSI6	0,146	0,021	0,102	0,306	-0,395	0,037	0,082	-0,128*	-0,024	0,050	-0,111	-0,018	0,070	-0,032	1	0,005	-0,018	-0,220
VMSI7	-0,012	-0,020	0,051	-0,033	0,043	-0,020	-0,021	0,060	0,795	0,004	-0,001	0,794	-0,065	-0,004	0,005	1	0,033	0,078
VMSI8	-0,006	-0,018	0,053	-0,022	0,066	-0,008	-0,068	0,172	0,030	0,013	-0,007	0,009	0,022	-0,018	-0,018	0,033	1	0,100
INSOLV	-0,204	-0,181	0,083	-0,310	0,069	-0,180	-0,002	0,096	0,107	-0,151	0,049	0,059	-0,166	0,107	-0,220	0,078	0,100	1

Como resumen del análisis exploratorio realizado, es posible concluir que de las diecisiete variables consideradas, prácticamente todas muestran signos de ser variables relevantes en el análisis de la solvencia.

4.2.3.2 Análisis Confirmatorio

4.2.3.2.1 Modelo Logit Sector Industria

En la tabla 4.16 se muestran los resultados del modelo Logit estimado para la muestra de empresas del sector industrial, un año antes de la quiebra (M.I), y que ha resultado ser la regresión logística que mejores resultados presentaba en lo que respecta a nivel de predicción. Las variables seleccionadas por el modelo han sido ocho: VMG1, VMG2, VMG3, VMG4, VMG5, VMG6, VMG8 y VMSI7.

Al analizar los signos de los coeficientes obtenidos para cada una de estas variables se puede apreciar que las variables VMG1, VMG2, VMG5 y VMG8 presentan signo negativo, lo cual denota una relación inversa con la insolvencia.

Por su parte, las variables VMG3, VMG4, VMG6 y VMSI7 presentan todas signos positivos. La relación con la insolvencia es directa, es decir, cuanto más alto el ratio mayor es la probabilidad de insolvencia. Además, las variables con signo negativo (VMG1, VMG2, VMG5 y VMG8) presentan un Odds ratio entre 0 y 1. Tan sólo es

superior a 1 en las variables que presentan signo positivo (VMG3, VMG4, VMG6 y VMSI7).

**TABLA 4.16 MODELO LOGIT
SECTOR INDUSTRIAL. M.I**

Ajuste del modelo		Valor	
Significatividad prueba Ómnibus		0,000	
Significatividad Test de Hosmer y Lemeshow		0,353	
-2 log verosimilitud (-2LL)		159,940	
R ² Cox & Snell		0,484	
R ² Nagelkerke		0,657	
Matriz de clasificación		%	
Dentro de la muestra		86,6	
		Insolventes	75,0
		Solventes	93,8
Muestra de testeo		89,2	
		Insolventes	82,0
		Solventes	95,1
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-39,344	0,000	0,000
VMG2	-1,422	0,241	0,000
VMG3	8,394	4.419,803	0,000
VMG4	16,676	17.477.639,31	0,027
VMG5	-1,371	0,254	0,001
VMG6	1,803	6,067	0,000
VMG8	-3,367	0,034	0,050
VMSI7	1,059	2,885	0,000
Constante	1,202	3,325	0,346

NOTA: VMG1 = Beneficio después de Impuestos/Total Activo; VMG2 = Activo Corriente/Pasivo Corriente; VMG3 = Fondo de Maniobra/Total Activo; VMG4 = EBIT/Total Activo; VMG5 = Total Ingresos/Total Activo; VMG6 = Quick Ratio; VMG8 = Activo Corriente/Total Activo; VMSI7 = Resultado de Explotación/Patrimonio Neto.

En el caso de VMG6 al tener signo positivo y un Odds ratio de 6,067 significa que cuando el ratio aumenta en un punto, la probabilidad de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 6,067 veces.

A continuación procede analizar distintos tests para constatar la robustez y precisión del modelo en su conjunto.

Para la prueba Ómnibus el valor obtenido es de 0,000, concluyéndose que el ajuste es adecuado, dado que al menos uno de los predictores está significativamente relacionado con la variable dependiente.

La matriz de clasificación, que evalúa la capacidad predictiva del modelo, muestra que se han clasificado correctamente el 86,6% de las empresas, obteniéndose un mayor nivel de predicción en el caso de las empresas solventes, con un 93,8% de acierto, y clasificando correctamente el 75,0% de las empresas insolventes. Por su parte, el nivel de clasificación en la muestra externa, utilizada para comprobar la capacidad de generalización del modelo, ha resultado ser del 89,2%.

Para el modelo del sector industrial un año antes de la quiebra, la significatividad del estadístico de Hosmer y Lemeshow es 0,353, lo que supondría aceptar la hipótesis de que no existen diferencias significativas entre las clasificaciones observadas y predichas.

El test $-2 \log$ de verosimilitud ($-2LL$) toma un valor de 159,940, lo que implica un buen ajuste del modelo.

Por último, el coeficiente de Cox y Snell ha tomado un valor de 0,484. Y por su parte, el R^2 de Nagelkerke alcanza un valor de 0,657. Ambos resultados denotan un ajuste aceptable del modelo seleccionado.

4.2.4 RESULTADOS MODELO CONSTRUCCIÓN

4.2.4.1 Análisis Exploratorio

En este epígrafe se aborda el análisis exploratorio para la muestra de empresas pertenecientes al sector de la construcción un año antes de la quiebra.

Se analizarán los principales parámetros estadísticos de las variables utilizadas (media, mediana, desviación típica, mínimo y máximo), para posteriormente analizar la normalidad de las mismas.

4.2.4.1.1 Análisis Descriptivo

Del análisis descriptivo realizado para las empresas que componen la muestra del sector de la construcción, un año antes de la entrada en quiebra, se ha podido constatar que tan solo VMG1, VMG4 y VMSC2 presentan medias con diferente signo, según se trate de empresas solventes e insolventes (tabla 4.17). Además, que el 50% de los valores medios en las empresas solventes supera al de las empresas insolventes. Para el caso de la mediana, este

porcentaje es del 50%, igualmente. En concreto, las variables VMG2, VMG3, VMG7, VMG8, VMG9 y VMSC1 presentan medias inferiores en el caso de las empresas solventes frente a las insolventes.

Asimismo, de la observación de las desviaciones típicas, así como de los mínimos y máximos, se comprueba la gran dispersión que existe en la mayor parte de los casos. Con respecto a la desviación típica, esta presenta valores aceptables, salvo para VMG2, VMG6 y VMSC3.

**TABLA 4.17 ESTADÍSTICOS DESCRIPTIVOS
(MODELO CONSTRUCCIÓN, M.I)**

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,05	-0,07	0,01	-0,01	0,10	0,22	-0,08	-1,90	0,59	0,30
	VMG2	4,04	4,20	1,69	1,66	7,74	7,69	0,00	0,23	55,70	53,46
	VMG3	0,33	0,61	0,28	0,66	0,30	0,32	-0,39	-0,68	0,97	0,99
	VMG4	-0,00	-0,04	0,00	0,00	0,43	0,21	-2,14	-1,85	2,02	0,35
	VMG5	0,72	0,65	0,34	0,43	0,89	0,77	0,00	0,00	4,65	4,56
	VMG6	2,27	0,83	0,92	0,52	4,44	1,60	0,00	0,00	25,66	14,13
	VMG7	0,34	0,55	0,29	0,59	0,29	0,30	0,00	0,00	1,92	1,49
	VMG8	0,60	0,81	0,67	0,88	0,31	0,22	0,00	0,05	1,00	1,00
	VMG9	0,24	0,35	0,04	0,03	1,12	2,05	-4,45	-8,68	7,91	13,94
VARIABLES MODELO SECTORIAL (CONSTRUCCIÓN)	VMSC1	0,03	0,14	0,00	0,04	0,71	1,19	-4,65	-3,83	1,69	4,16
	VMSC2	0,34	-0,19	0,09	0,01	0,80	1,26	-2,12	-9,98	5,60	3,46
	VMSC3	5,11	0,36	1,10	0,08	10,31	2,77	-0,77	-11,53	52,81	20,51

4.2.4.1.2 Análisis de Normalidad

En la tabla 4.18 se muestran los resultados del test de Kolmogorov-Smirnov para las empresas solventes e insolventes pertenecientes al sector construcción, a un nivel de significación del 5%. Se observa cómo todas las variables tienen un nivel de significación inferior al 5%, lo que implica rechazar la hipótesis nula de normalidad. Esto supone que, para proseguir con el análisis descriptivo, deberán aplicarse test no paramétricos.

**TABLA 4.18 ANÁLISIS DE NORMALIDAD
(MODELO CONSTRUCCIÓN, M.I)**

		M.I (1 AÑO ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	0,297	0,000	0,269	0,000
	VMG2	0,452	0,000	0,302	0,000
	VMG3	0,076	0,005	0,124	0,000
	VMG4	0,335	0,000	0,347	0,000
	VMG5	0,322	0,000	0,217	0,000
	VMG6	0,462	0,000	0,303	0,000
	VMG7	0,213	0,000	0,075	0,006
	VMG8	0,148	0,000	0,203	0,000
	VMG9	0,325	0,000	0,218	0,000
VARIABLES MODELO SECTORIAL (CONSTRUCCIÓN)	VMSC1	0,265	0,000	0,127	0,000
	VMSC2	0,359	0,000	0,301	0,000
	VMSC3	0,309	0,000	0,341	0,000

NOTA: K-S: test de Kolmogorov-Smirnov

4.2.4.1.3 Tests no paramétricos

Los resultados de la aplicación de los tests correspondientes aparecen en la tabla 4.19. Las variables VMG2 y VMG5 muestran en el test de Mann-Whitney y de Kolmogorov-Smirnov un nivel de significación superior al 5% ($p > 0,05$), con lo cual se aceptaría solo en estos casos la hipótesis nula de igualdad de distribuciones. Asimismo, VMG9 y VMSC1 presentan tests de Mann-Whitney con un nivel de significación superior al 5%. Por tanto, con las excepciones apuntadas, se rechaza la hipótesis nula de igualdad en las distribuciones en todas las variables consideradas en el sector construcción.

**TABLA 4.19 TESTS NO PARAMÉTRICOS
(MODELO CONSTRUCCIÓN, M.I)**

		M.I (1 AÑO ANTES)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	-12,266	0,000	6,364	0,000
	VMG2	-0,602	0,547	0,808	0,532
	VMG3	8,088	0,000	3,978	0,000
	VMG4	-3,254	0,001	2,119	0,000
	VMG5	-1,878	0,060	1,312	0,064
	VMG6	-6,389	0,000	3,233	0,000
	VMG7	9,064	0,000	4,245	0,000
	VMG8	4,656	0,000	2,733	0,000
	VMG9	0,102	0,919	2,527	0,000
VARIABLES MODELO SECTORIAL (CONSTRUCCIÓN)	VMSC1	-1,158	0,247	2,216	0,000
	VMSC2	-9,358	0,000	4,903	0,000
	VMSC3	-9,445	0,000	4,581	0,000

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

4.2.4.1.4 Análisis de Correlación

En la tabla 4.20 se muestra el análisis de correlación bilateral para la muestra del sector construcción, un año antes de la quiebra. Puede comprobarse cómo solo se registra una elevada correlación entre los pares de variables VMG2 y VMG6 (0,996).

**Tabla 4.20 ANÁLISIS DE CORRELACIÓN BILATERAL
(SECTOR CONSTRUCCIÓN, M.I)**

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	VMSC1	VMSC2	VMSC3	INSOLV
VMG1	1	0,035	0,029	0,107	0,396	0,035	-0,250	-0,009	-0,118	0,082	0,144	0,229	-0,353
VMG2	0,035	1	-0,013	0,015	-0,030	0,996	-0,030	0,052	-0,009	0,006	0,020	0,147	-0,056
VMG3	0,029	-0,013	1	0,023	-0,142	-0,034	0,371	0,570	-0,039	0,004	-0,052	-0,196	0,367
VMG4	0,107	0,015	0,023	1	-0,066	0,011	-0,010	-0,076	-0,049	0,067	0,094	0,139	-0,040
VMG5	0,396	-0,030	-0,142	-0,066	1	-0,015	-0,243	0,075	-0,009	0,023	-0,003	-0,028	-0,150
VMG6	0,035	0,996	-0,034	0,011	-0,015	1	-0,056	0,043	-0,004	0,004	0,017	0,234	-0,062
VMG7	-0,250	-0,030	0,371	-0,010	-0,243	-0,056	1	0,060	-0,063	-0,036	-0,091	-0,365	0,388
VMG8	-0,009	0,052	0,570	-0,076	0,075	0,043	0,060	1	0,005	0,035	-0,061	-0,119	0,259
VMG9	-0,118	-0,009	-0,039	-0,049	-0,009	-0,004	-0,063	0,005	1	0,025	-0,044	-0,024	0,042
VMSC1	0,082	0,006	0,004	0,067	0,023	0,004	-0,036	0,035	0,025	1	0,075	-0,030	0,003
VMSC2	0,144	0,020	-0,052	0,094	-0,003	0,017	-0,091	-0,061	-0,044	0,075	1	0,233	-0,229
VMSC3	0,229	0,147	-0,196	0,139	-0,028	0,234	-0,365	-0,119	-0,024	-0,030	0,233	1	-0,326
INSOLV	-0,353	-0,056	0,367	-0,040	-0,150	-0,062	0,388	0,259	0,042	0,003	-0,229	-0,326	1

4.2.4.2 Análisis Confirmatorio

En este apartado se estimará un modelo Logit para la muestra seleccionada del sector construcción, con información de un año antes de producirse la quiebra.

El modelo obtenido se testará en una muestra externa, lo que permitirá validar la capacidad de generalización del mismo. Para ello, la muestra se dividirá en dos partes: un 70% para la denominada muestra de entrenamiento y un 30% para la denominada muestra de testeo.

4.2.4.2.1 Modelo Logit Sector Construcción

En la tabla 4.21 aparecen los resultados del Logit estimado para la muestra de empresas del sector construcción, un año antes de la quiebra. Las variables seleccionadas en el modelo han sido VMG1, VMG6, VMG7, VMG8 y VMSC1.

Al analizar los signos de los coeficientes obtenidos para cada una de estas variables, se comprueba que VMG1 y VMG6 presentan signo negativo, lo cual denota una relación inversa con la insolvencia. En consecuencia, cuanto más altos son estos ratios menor es la probabilidad de que la empresa entre en quiebra.

Analizando los Odds ratio de las variables, como cabía esperar, observamos que VMG1 y VMG6, al presentar signo negativo, tienen valores entre 0 y 1. Son

mayores que 1 las variables con signo positivo (VMG7, VMG8 y VMSC1).

**TABLA 4.21 MODELO LOGIT
SECTOR CONSTRUCCIÓN. M.I**

Ajuste del modelo		Valor	
Significatividad prueba Ómnibus		0,000	
Significatividad Test de Hosmer y Lemeshow		0,177	
-2 log verosimilitud (-2LL)		260,006	
R ² Cox & Snell		0,417	
R ² Nagelkerke		0,559	
Matriz de clasificación		%	
Dentro de la muestra		80,8	
	Insolventes	85,1	
	Solventes	75,4	
Muestra de testeo		81,5	
	Insolventes	84,1	
	Solventes	78,6	
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-14,624	0,000	0,000
VMG6	-0,138	0,871	0,036
VMG7	1,377	3,964	0,054
VMG8	2,134	8,448	0,018
VMSC1	0,390	1,477	0,034
Constante	-2,092	0,123	0,001

NOTA: VMG1 = Beneficio después de Impuestos/Total Activo; VMG6 = Quick Ratio; VMG7 = Total Deuda/Total Activo; VMG8 = A corriente/Total Activo; VMSC1 = Gastos Financieros/EBIT

El caso de VMG7, al tener signo positivo y un Odds ratio de 3,964, significa que cuando el ratio Total Deuda/Total Activo aumenta en un punto, la probabilidad

de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 3,964 veces.

Con respecto a la significatividad de los coeficientes del modelo, puede apreciarse que todos ellos son significativos con un p-valor inferior o igual al 5%.

Una vez valorada la significación de cada variable de forma individual, resulta conveniente analizar distintos tests para constatar la robustez y precisión del modelo en su conjunto.

En primer lugar, la prueba de Razón de Verosimilitud (RV) de todos los coeficientes, o también denominada prueba Ómnibus, estudia la significatividad conjunta de las variables seleccionadas. En el caso del modelo obtenido con la muestra de empresas del sector de la construcción en M.I, el valor obtenido es de 0,000 concluyéndose que el ajuste es adecuado, dado que al menos uno de los predictores está significativamente relacionado con la variable dependiente.

En segundo lugar, utilizando la matriz de clasificación, se observa que se han clasificado correctamente el 80,8% de las empresas, obteniéndose un mayor nivel de predicción en el caso de las empresas insolventes, con un 85,1% de acierto, y clasificando correctamente el 75,4% de las empresas solventes.

Por su parte, el nivel de clasificación en la muestra externa, utilizada para comprobar la capacidad de generalización del modelo, ha resultado ser del 81,5%.

En el test de Hosmer y Lemeshow la significatividad del estadístico es 0,177 y, en consecuencia, se aceptaría la hipótesis de que no existen diferencias significativas entre las clasificaciones observadas y predichas.

El test $-2 \log$ de verosimilitud ($-2LL$), como medida global del ajuste del modelo, toma un valor, para la M.I, de 260,006.

Igualmente, el coeficiente de Cox y Snell, toma un valor para M.I de 0,417. El R^2 de Nagelkerke es una modificación del coeficiente de Cox y Snell, y alcanza un valor de 0,559. Ambos resultados denotan un ajuste aceptable del modelo seleccionado.

4.2.5 RESULTADOS MODELO COMERCIO Y SERVICIOS

4.2.5.1 Análisis Exploratorio

En este apartado se aborda el análisis exploratorio para la muestra de empresas pertenecientes al sector comercio y servicios.

4.2.5.1.1 Análisis Descriptivo

Conforme a los resultados del análisis descriptivo realizado (tabla 4.22), se comprueba que sólo VMG1, VMG4, VMSCS1, VMSCS2 y VMSCS3 presentan medias con diferente signo según se trate de empresas solventes e insolventes. Además, en la mayor parte de los casos los valores son superiores para las empresas solventes (el 64% de los valores medios de las variables correspondientes a empresas solventes supera al de las empresas insolventes).

**TABLA 4.22 ESTADÍSTICOS DESCRIPTIVOS
(MODELO COMERCIO Y SERVICIOS, M.I)**

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,02	-0,14	0,01	-0,05	0,05	0,22	-0,15	-1,19	0,28	0,13
	VMG2	1,52	1,07	1,20	0,97	1,01	0,69	0,19	0,11	7,60	5,41
	VMG3	0,28	0,35	0,28	0,30	0,29	0,31	-0,39	-1,08	0,92	0,96
	VMG4	0,04	-0,12	0,03	-0,03	0,06	0,23	-0,14	-1,14	0,39	0,21
	VMG5	1,66	1,19	1,33	0,93	1,26	0,90	0,10	0,00	7,30	4,91
	VMG6	0,94	0,59	0,74	0,51	0,93	0,49	0,00	0,00	6,08	4,24
	VMG7	0,21	0,43	0,15	0,42	0,21	0,30	0,00	0,00	0,81	1,59
	VMG8	0,70	0,61	0,76	0,63	0,24	0,23	0,03	0,02	1,00	0,99
	VMG9	0,10	0,18	0,05	0,05	0,61	1,19	-4,14	-3,84	4,40	4,69
VARIABLES MODELO SECTORIAL (COMERCIO Y SERVICIOS)	VMSCS1	0,07	-0,09	0,05	0,00	0,08	0,23	-0,14	-1,12	0,51	0,21
	VMSCS2	11,33	-6,65	2,59	-0,89	37,25	16,79	-86,75	-102,81	340,00	19,01
	VMSCS3	0,12	-0,13	0,07	-0,05	0,21	0,34	-0,26	-1,59	1,46	1,41
	VMSCS4	21,22	32,78	6,60	4,73	61,04	135,17	0,20	0,00	550,99	1468,42
	VMSCS5	1,69	1,19	1,33	0,93	1,26	0,90	0,10	0,00	7,30	4,91

Sin embargo, las variables VMG3, VMG7, VMG9 y VMSCS4 presentan medias inferiores en el caso de las empresas solventes.

Por su parte, la desviación típica presenta valores moderados, salvo para VMSCS2 y VMSCS4 en M.I.

4.2.5.1.2 Análisis de Normalidad

En la tabla 4.23 se muestran los resultados del test de Kolmogorov-Smirnov para las empresas solventes e insolventes pertenecientes al sector comercio y servicios, a un nivel de significación del 5%.

**TABLA 4.23 ANÁLISIS DE NORMALIDAD
(MODELO COMERCIO Y SERVICIOS, M.I)**

		M.I (1 AÑO ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO SECTORIAL (COMERCIO Y SERVICIOS)	VMG1	0,228	0,000	0,252	0,000
	VMG2	0,236	0,000	0,446	0,000
	VMG3	0,056	0,200	0,118	0,000
	VMG4	0,192	0,000	0,219	0,000
	VMG5	0,212	0,000	0,160	0,000
	VMG6	0,297	0,000	0,448	0,000
	VMG7	0,213	0,000	0,195	0,000
	VMG8	0,124	0,000	0,085	0,001
	VMG9	0,420	0,000	0,159	0,000
VARIABLES MODELO SECTORIAL (COMERCIO Y SERVICIOS)	VMSCS1	0,155	0,000	0,189	0,000
	VMSCS2	0,313	0,000	0,510	0,000
	VMSCS3	0,270	0,000	0,493	0,000
	VMSCS4	0,371	0,000	0,399	0,000
	VMSCS5	0,212	0,000	0,168	0,000

Se aprecia cómo, con la excepción de VMG3 y para empresas solventes e insolventes, todas las variables en M.I tienen un nivel de significación inferior al 5%. Esto implica rechazar la hipótesis nula de normalidad.

4.2.5.1.3 Tests no paramétricos

La aplicación de los correspondientes tests no paramétricos han permitido comprobar que tan solo VMG3 y VMSCS4 registran un nivel de significación superior al 5% ($p > 0,05$), con lo cual se aceptaría sólo en estos casos la hipótesis nula de igualdad de distribuciones (tabla 4.24).

**TABLA 4.24 TESTS NO PARAMÉTRICOS
(MODELO COMERCIO Y SERVICIOS, M.I)**

		M.I (1 AÑO ANTES DE LA QUIEBRA)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO SECTORIAL (COMERCIO Y SERVICIOS)	VMG1	-11,653	0,000	5,513	0,000
	VMG2	-6,574	0,000	3,282	0,000
	VMG3	1,198	0,231	1,344	0,054
	VMG4	-9,325	0,000	4,927	0,000
	VMG5	-4,322	0,000	1,891	0,002
	VMG6	-4,898	0,000	2,756	0,000
	VMG7	8,548	0,000	3,956	0,000
	VMG8	-3,470	0,001	1,928	0,001
	VMG9	1,794	0,073	2,678	0,000
VARIABLES MODELO SECTORIAL (COMERCIO Y SERVICIOS)	VMSCS1	-9,356	0,000	4,763	0,000
	VMSCS2	-11,282	0,000	5,800	0,000
	VMSCS3	-9,739	0,000	4,936	0,000
	VMSCS4	-1,619	0,105	1,329	0,059
	VMSCS5	-4,318	0,000	1,891	0,002

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

Con la excepción apuntada, se rechazaría la hipótesis nula de igualdad en las distribuciones en todas las variables consideradas en el sector comercio y servicios.

A partir de los resultados obtenidos, y dado el gran porcentaje de significación de casi la totalidad de las variables, parece posible la obtención de adecuados modelos de predicción, por tratarse de variables que, a priori, resultan ser apropiadas para valorar la insolvencia.

4.2.5.1.4 Análisis de Correlación

Del análisis de correlación efectuado (tabla 4.25) se deduce la existencia de una elevada correlación entre los pares de variables VMG1 y VMG4 (0,989), VMG2 y VMG6 (0,997), VMG2 y VMSCS3 (-0,985), VMG4 y VMSCS1 (0,984), VMG5 y VMSCS5 (1,000), VMSCS1 y VMG1 (0,964) y VMSCS3 y VMG6 (-0,968).

TABLA 4.25 ANÁLISIS DE CORRELACIÓN BILATERAL
(MODELO COMERCIO Y SERVICIOS, M.I)

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	VMSCS1	VMSCS2	VMSCS3	VMSCS4	VMSCS5	INSOLV
VMG1	1	-0,214	0,319	0,989	-0,122	-0,208	-0,224	-0,042	-0,091	0,964	0,031	0,264	-0,003	-0,128	-0,342
VMG2	-0,214	1	0,074	0,194	-0,051	0,997	-0,133	0,094	0,000	-0,255	0,074	-0,985	-0,019	-0,050	0,011
VMG3	0,319	0,074	1	0,323	-0,137	0,047	-0,150	0,342	-0,040	0,103	-0,032	-0,049	-0,060	-0,125	0,012
VMG4	0,989	0,194	0,323	1	-0,121	0,130	-0,276	0,009	-0,089	0,984	0,037	0,494	-0,002	-0,125	-0,340
VMG5	-0,122	-0,051	-0,137	-0,121	1	-0,047	-0,192	0,185	0,022	0,068	0,002	0,041	0,228	1,000	-0,182
VMG6	-0,208	0,997	0,047	0,130	-0,047	1	-0,121	0,095	0,003	-0,249	0,057	-0,968	-0,006	-0,047	0,003
VMG7	-0,224	-0,133	-0,150	-0,276	-0,192	-0,121	1	-0,251	0,010	-0,284	-0,048	-0,327	-0,084	-0,193	0,340
VMG8	-0,042	0,094	0,342	0,009	0,185	0,095	-0,251	1	0,071	-0,025	-0,055	-0,060	-0,019	0,191	-0,167
VMG9	-0,091	0,000	-0,040	-0,089	0,022	0,003	0,010	0,071	1	-0,097	-0,006	-0,004	0,013	0,026	0,066
VMSCS1	0,964	-0,255	0,103	0,984	0,068	-0,249	-0,284	-0,025	-0,097	1	0,037	0,305	-0,004	0,059	-0,365
VMSCS2	0,031	0,074	-0,032	0,037	0,002	0,057	-0,048	-0,055	-0,006	0,037	1	0,102	0,015	0,001	0,040
VMSCS3	0,264	-0,985	-0,049	0,494	0,041	-0,968	-0,327	-0,060	-0,004	0,305	0,102	1	0,015	0,040	-0,060
VMSCS4	-0,003	-0,019	-0,060	-0,002	0,228	-0,006	-0,084	-0,019	0,013	-0,004	0,015	0,015	1	0,220	0,036
VMSCS5	-0,128	-0,050	-0,125	-0,125	1,000	-0,047	-0,193	0,191	0,026	0,059	0,001	0,040	0,220	1	-0,178
INSOLV	-0,342	0,011	0,012	-0,340	-0,182	0,003	0,340	-0,167	0,066	-0,365	0,040	-0,060	0,036	-0,178	1

4.2.5.2 Análisis Confirmatorio

En este apartado se presentan los resultados del modelo Logit para la muestra del sector comercio y servicios, con información de un año antes de producirse la quiebra.

4.2.5.2.1 Modelo Logit sector comercio y servicios

En la tabla 4.26 se muestran los resultados del modelo estimado para la muestra de empresas del sector comercio y servicios, un año antes de la quiebra, M.I. El modelo estimado para la muestra de empresas del sector comercio y servicios está formado por las variables VMG1, VMG2, VMG3, VMG7, VMG8 y VMSCS4.

Con respecto al signo de los coeficientes obtenidos para cada una de estas variables, se puede apreciar que las variables VMG1 y VMG8 presentan signo negativo, lo cual denota una relación inversa con la insolvencia. Por su parte, las variables VMG3, VMG7 y VMSCS4 presentan signos positivos. La relación con la insolvencia en este caso es directa, es decir, cuanto más alto es el ratio mayor es la probabilidad de insolvencia.

Respecto a VMG7, al tener signo positivo y un Odds ratio de 8,715, se deduce que cuando el ratio Total Deuda/Total Activo aumenta en un punto, la probabilidad de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 8,715 veces.

Asimismo, se comprueba que todas las variables incluidas en el modelo son significativas al nivel del 10%.

TABLA 4.26 MODELO LOGIT
SECTOR COMERCIO Y SERVICIOS. M.I

Ajuste del modelo		Valor	
Significatividad prueba Ómnibus		0,000	
Significatividad Test de Hosmer y Lemeshow		0,149	
-2 log verosimilitud (-2LL)		217,607	
R ² Cox & Snell		0,438	
R ² Nagelkerke		0,584	
Matriz de clasificación		%	
Dentro de la muestra		79,9	
Insolventes		76,3	
Solventes		83,8	
Muestra de testeo		83,5	
Insolventes		86,5	
Solventes		79,5	
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-28,061	0,000	0,000
VMG2	-0,785	0,456	0,024
VMG3	2,992	19,932	0,001
VMG7	2,165	8,715	0,010
VMG8	-2,112	0,121	0,056
VMSCS4	0,006	1,006	0,040
Constante	-0,066	0,936	0,936

Notas: VMG1 = Beneficio después de Impuestos/Total Activo; VMG2 = Activo Corriente/Pasivo Corriente; VMG3 = Fondo de Maniobra/Total Activo; VMG7 = Total Deuda/Total Activo; VMG8 = Activo Corriente/Total Activo; VMSCS4 = Ventas/Existencias

Una vez valorada la significación de cada variable de forma individual, procede analizar distintos tests para constatar la robustez y precisión del modelo en su conjunto.

El valor obtenido en la prueba de Razón de Verosimilitud (RV) es de 0,000 por lo que el ajuste es adecuado.

La matriz de clasificación muestra cómo se han clasificado correctamente el 79,9% de las empresas, obteniéndose un mayor nivel de predicción en el caso de las empresas solventes, con un 83,8% de acierto, y clasificando correctamente el 76,3% de las empresas insolventes.

Por su parte, el nivel de clasificación en la muestra externa, utilizada para comprobar la capacidad de generalización del modelo, ha resultado ser del 83,5%.

La bondad del ajuste del modelo, a tenor de lo que señala el test de Hosmer y Lemeshow, muestra una significatividad del 0,149 por lo que se aceptaría la hipótesis de que no existen diferencias significativas entre las clasificaciones observadas y predichas.

El test $-2 \log$ de verosimilitud ($-2LL$), como medida de ajuste global del modelo, apunta un valor de 217,607.

Por su parte, el coeficiente de Cox y Snell y el R^2 de Nagelkerke muestran valores de 0,438 y 0,584,

respectivamente. Ambos resultados denotan un ajuste aceptable del modelo seleccionado.

4.2.6 RESULTADOS SECTOR HOSTELERÍA

4.2.6.1 Análisis Exploratorio

En este epígrafe se aborda el análisis exploratorio para la muestra de empresas pertenecientes al sector hostelería.

Al igual que en los sectores anteriores, se analizarán los principales parámetros estadísticos de las variables utilizadas (media, mediana, desviación típica, mínimo y máximo), para posteriormente analizar la normalidad y la correlación entre las variables.

4.2.6.1.1 Análisis Descriptivo

El análisis descriptivo realizado ha permitido comprobar que VMG1, VMG4, VMSH1, VMSH2, VMSH3 y VMSH6 presentan medias con diferente signo (tabla 4.27).

Además, que tan solo VMG7 presenta una media inferior en el caso de las empresas solventes frente a las insolventes. Y con respecto a la desviación típica, que ésta presenta valores moderados, salvo para VMSH5.

**TABLA 4.27 ESTADÍSTICOS DESCRIPTIVOS
(MODELO HOSTELERÍA, M.I)**

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,03	-0,16	0,02	-0,08	0,06	0,22	-0,44	-0,92	0,26	0,08
	VMG2	1,61	0,68	1,07	0,44	1,73	0,73	0,00	0,02	8,36	3,68
	VMG3	0,13	0,06	0,07	0,04	0,22	0,20	-0,38	-0,38	0,79	0,74
	VMG4	0,06	-0,15	0,04	-0,07	0,07	0,22	-0,16	-0,87	0,35	-0,12
	VMG5	1,95	1,25	1,59	0,89	1,52	1,21	0,00	0,00	7,39	5,84
	VMG6	1,14	0,54	0,64	0,28	1,49	0,66	0,00	0,00	7,71	3,45
	VMG7	0,42	0,56	0,36	0,53	0,35	0,35	0,00	0,00	2,33	1,67
	VMG8	0,40	0,32	0,36	0,26	0,27	0,26	0,00	0,00	0,99	1,00
	VMG9	0,21	-0,00	0,08	0,08	0,50	1,54	-0,67	-9,91	3,14	4,23
VARIABLES MODELO SECTORIAL (HOSTELERÍA)	VMSH1	0,49	-0,13	0,37	-0,09	0,51	0,43	-0,48	-2,13	3,08	1,03
	VMSH2	0,11	-0,10	0,09	-0,03	0,08	0,21	-0,16	-0,85	0,40	0,23
	VMSH3	6,13	-5,37	2,99	-1,26	9,90	23,29	-25,23	-105,66	51,05	61,64
	VMSH4	3,71	1,45	0,84	0,08	8,59	9,80	-8,35	-47,34	44,74	32,33
	VMSH5	250,16	152,45	38,95	12,12	806,83	608,31	0,12	0,01	6435,7	5357,0
	VMSH6	1,16	-0,12	0,21	-0,08	3,50	2,32	-0,18	-12,61	22,19	19,53

4.2.6.1.2 Análisis de Normalidad

En la tabla 4.28 se muestran los resultados del test de Kolmogorov-Smirnov para las empresas solventes e insolventes pertenecientes al sector hostelería, a un nivel de significación del 5%.

Todas las variables presentan un nivel de significación inferior al 5%, lo que implica rechazar la hipótesis nula de normalidad.

**TABLA 4.28 ANÁLISIS DE NORMALIDAD
(MODELO HOSTELERÍA, M.I)**

		M.I (1 AÑO ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	0,369	0,000	0,341	0,000
	VMG2	0,408	0,000	0,342	0,000
	VMG3	0,143	0,000	0,134	0,000
	VMG4	0,329	0,000	0,334	0,000
	VMG5	0,427	0,000	0,279	0,000
	VMG6	0,419	0,000	0,218	0,000
	VMG7	0,220	0,000	0,136	0,000
	VMG8	0,090	0,000	0,124	0,000
	VMG9	0,345	0,000	0,164	0,000
VARIABLES MODELO SECTORIAL (HOSTELERÍA)	VMSH1	0,381	0,000	0,186	0,000
	VMSH2	0,303	0,000	0,304	0,000
	VMSH3	0,340	0,000	0,272	0,000
	VMSH4	0,401	0,000	0,255	0,000
	VMSH5	0,512	0,000	0,519	0,00
	VMSH6	0,367	0,000	0,353	0,000

NOTA: K-S: test de Kolmogorov-Smirnov

4.2.6.1.3 Tests no paramétricos

Los resultados que aparecen en la tabla 4.29 indican que tan sólo VMG9 presenta un test de Mann-Whitney con un nivel de significación superior al 5% ($p > 0,05$), con lo cual se aceptaría sólo en este caso la hipótesis nula de igualdad de distribuciones. Con la excepción apuntada, se rechaza la hipótesis nula de igualdad en las distribuciones mediante los resultados obtenidos por la aplicación del test de Mann-Whitney y de Kolmogorov-Smirnov en todas las variables consideradas en el sector hostelería.

**TABLA 4.29 TESTS NO PARAMÉTRICOS
(MODELO HOSTELERÍA, M.I)**

		M.I (1 AÑO ANTES)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	-14,357	0,000	7,534	0,000
	VMG2	-8,936	0,000	4,388	0,000
	VMG3	-3,381	0,001	1,785	0,003
	VMG4	-13,347	0,000	7,050	0,000
	VMG5	-5,688	0,000	2,673	0,000
	VMG6	-6,617	0,000	3,216	0,000
	VMG7	4,579	0,000	2,309	0,000
	VMG8	-4,847	0,000	2,228	0,000
	VMG9	0,831	0,406	2,370	0,000
VARIABLES MODELO SECTORIAL (HOSTELERÍA)	VMSH1	-13,406	0,000	6,429	0,000
	VMSH2	-12,701	0,000	6,476	0,000
	VMSH3	-9,087	0,000	5,270	0,000
	VMSH4	-4,063	0,000	3,153	0,000
	VMSH5	-6,444	0,000	3,403	0,000
	VMSH6	-11,123	0,000	5,876	0,000

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

4.2.6.1.4 Análisis de Correlación

Conforme a los resultados del coeficiente de correlación de Pearson (tabla 4.30), se ha podido comprobar una elevada correlación entre los pares de variables VMG1 y VMG4 (0,996), VMG2 y VMG6 (0,998), VMG1 y VMSH2 (0,990) y VMG4 y VMSH2 (0,996).

**TABLA 4.30 ANÁLISIS DE CORRELACIÓN BILATERAL
(MODELO HOSTELERÍA, M.I)**

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	VMSH1	VMSH2	VMSH3	VMSH4	VMSH5	VMSH6	INSOLV
VMG1	1	0,052	0,004	0,996	0,234	0,049	-0,005	0,025	-0,134	0,088	0,990	0,031	0,019	0,011	0,261	-0,257
VMG2	0,052	1	0,068	0,039	-0,011	0,988	-0,066	0,135	0,005	0,462	0,037	0,013	-0,011	-0,010	0,081	-0,139
VMG3	0,004	0,068	1	0,002	-0,013	0,034	-0,117	0,403	-0,069	0,016	-0,001	-0,135	-0,018	-0,041	0,062	-0,106
VMG4	0,996	0,039	0,002	1	0,235	0,035	-0,001	0,047	-0,125	0,109	0,996	0,032	0,019	0,009	0,248	-0,265
VMG5	0,234	-0,011	-0,013	0,235	1	-0,011	-0,086	0,150	-0,147	0,039	0,259	-0,004	-0,014	0,192	0,175	-0,076
VMG6	0,049	0,988	0,034	0,035	-0,011	1	-0,097	0,115	0,005	0,434	0,032	-0,004	-0,024	-0,010	0,198	-0,126
VMG7	-0,005	-0,066	-0,117	-0,001	-0,086	-0,097	1	-0,335	0,062	-0,078	-0,016	0,010	0,089	-0,063	-0,242	0,195
VMG8	0,025	0,135	0,403	0,047	0,150	0,115	-0,335	1	-0,003	0,116	0,040	0,007	-0,095	-0,015	0,159	-0,228
VMG9	-0,134	0,005	-0,069	-0,125	-0,147	0,005	0,062	-0,003	1	-0,025	-0,130	0,087	0,385	-0,013	0,017	0,005
VMSH1	0,088	0,462	0,016	0,109	0,039	0,434	-0,078	0,116	-0,025	1	0,111	0,029	0,004	-0,003	0,219	-0,182
VMSH2	0,990	0,037	-0,001	0,996	0,259	0,032	-0,016	0,040	-0,130	0,111	1	0,033	0,016	0,021	0,272	-0,271
VMSH3	0,031	0,013	-0,135	0,032	-0,004	-0,004	0,010	0,007	0,087	0,029	0,033	1	0,174	-0,002	0,003	-0,249
VMSH4	0,019	-0,011	-0,018	0,019	-0,014	-0,024	0,089	-0,095	0,385	0,004	0,016	0,174	1	-0,007	-0,018	-0,069
VMSH5	0,011	-0,010	-0,041	0,009	0,192	-0,010	-0,063	-0,015	-0,013	-0,003	0,021	-0,002	-0,007	1	0,047	-0,034
VMSH6	0,261	0,081	0,062	0,248	0,175	0,198	-0,242	0,159	0,017	0,219	0,272	0,003	-0,018	0,047	1	-0,210
INSOLV	-0,257	-0,139	-0,106	-0,265	-0,076	0,126	0,195	-0,228	0,005	-0,182	-0,271	-0,249	-0,069	-0,034	-0,210	1

4.2.6.2 Análisis Confirmatorio

En este apartado se realizará un análisis confirmatorio de los resultados obtenidos para el sector hostelería con datos de un año antes de producirse la quiebra.

4.2.6.2.1 Modelo Logit Sector Hostelería

Como en el resto de los modelos estimados, se aplicará el método de máximo verosimilitud, de forma que se irán realizando sucesivas iteraciones hasta obtener la estimación más probable.

En la tabla 4.31 se muestran los resultados del modelo estimado para empresas del sector hostelería, un año antes de la quiebra, M.I. Las variables seleccionadas por el modelo han sido VMG1, VMG2, VMG5, VMSH1 y VMSH3.

Al analizar los signos de los coeficientes obtenidos para cada una de estas variables, se puede apreciar que todas presentan signo negativo, lo cual denota una relación inversa con la insolvencia.

Observando los Odds ratio de las variables, como cabía esperar, al presentar todas las variables signo negativo, el Odds ratio muestra valores entre 0 y 1. Así, en el caso de VMSH3, al tener un Odds ratio de 0,973 indica que cuando el ratio Total Deuda Financiera/EBITDA aumenta en un punto, la probabilidad de que la empresa

sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 0,973 veces.

Por último, se ha podido constatar que todas las variables son significativas al nivel del 5%.

Respecto a la robustez del modelo, la prueba Ómnibus arroja una significación de 0,000, mostrando pues un ajuste adecuado. En la matriz de clasificación se observa que se han clasificado correctamente el 84,2% de las empresas, obteniéndose un mayor nivel de predicción en el caso de las empresas solventes, con un 89,2% de acierto, y clasificando correctamente el 79,0% de las empresas insolventes. Además, el nivel de clasificación en la muestra externa ha resultado ser del 91,2%.

Por su parte, la significatividad del estadístico del test de Hosmer y Lemeshow es 0,131 y, en consecuencia, se aceptaría la hipótesis de que no existen diferencias significativas entre las clasificaciones observadas y predichas.

Por otro lado, el test -2 log de verosimilitud (-2LL) toma un valor de 249,873, y los coeficientes de Cox y Snell y R^2 de Nagelkerke ascienden a 0,517 y 0,690, respectivamente.

Los resultados denotan, pues, un buen ajuste del modelo seleccionado.

TABLA 4.31 MODELO LOGIT
SECTOR HOSTELERÍA. M.I

Ajuste del modelo		Valor	
Significatividad prueba Ómnibus		0,000	
Significatividad Test de Hosmer y Lemeshow		0,131	
-2 log verosimilitud (-2LL)		249,873	
R ² Cox & Snell		0,517	
R ² Nagelkerke		0,690	
Matriz de clasificación		%	
Dentro de la muestra		84,2	
Insolventes		79,0	
Solventes		89,2	
Muestra de testeo		91,2	
Insolventes		91,5	
Solventes		90,9	
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-13,196	0,000	0,000
VMG2	-1,066	0,344	0,019
VMG5	-0,440	0,644	0,000
VMSH1	-1,447	0,235	0,007
VMSH3	-0,027	0,973	0,012
Constante	1,186	3,274	0,000

Notas: VMG1 = Beneficio después de Impuestos/Total Activo; VMG2 = Activo Corriente/Pasivo Corriente; VMG5 = Total Ingresos/Total Activo; VMSH1 = EBITDA/Pasivo Corriente; VMSH3 = Total Deuda Financiera/EBITDA.

4.2.7 COMPARACIÓN DEL MODELO GLOBAL CON LOS MODELOS SECTORIALES EN M.I.

Una vez obtenidos los modelos de regresión para cada uno de los sectores objeto de este estudio y para la muestra global compuesta por empresas de todos los sectores, procede contrastar la hipótesis H_1 formulada en el capítulo primero de esta tesis.

Esta hipótesis planteaba si un modelo global, obtenido con una muestra de empresas de los cinco sectores representados en este estudio, era igual o distinto a los modelos sectoriales estimados a partir de muestras de cada uno de los sectores incorporados en nuestro trabajo.

Para dar respuesta a esta hipótesis se utilizará los resultados que aparecen en la tabla 4.32. Como puede comprobarse, de la comparación de los modelos estimados es posible concluir que los conjuntos de variables seleccionados por los mismos son distintos. En consecuencia, la hipótesis H_1 es aceptada, en el sentido de que los modelos globales y sectoriales obtenidos para las muestras correspondientes a un año antes de la entrada en quiebra son diferentes.

TABLA 4.32 MODELOS GLOBAL Y SECTORIALES OBTENIDOS PARA M.I

Ajuste del Modelo	Global	Agricultura	Industria	Construcción	Comercio y Servicios	Hostelería
Prueba Ómnibus	0,000	0,000	0,000	0,000	0,000	0,000
Test Hosmer-Lemeshow	0,085	0,487	0,353	0,177	0,149	0,131
R ² Nagelkerke	0,441	0,560	0,657	0,559	0,584	0,690
Matriz de Clasificación (%)						
Fuera de la Muestra						
Solventes	86,9	57,9	95,1	78,6	79,5	90,9
Insolventes	74,5	90,9	82,0	84,1	86,5	91,5
Total	80,8	75,6	89,2	81,5	83,5	91,2
Predictores	VMG1 VMG3 VMG4 VMG5	VMG1 VMG6 VMSA2	VMG1	VMG1 VMG6 VMG7 VMG8 VMSC1	VMG1 VMG2 VMG3 VMG7 VMG8 VMSCS4	VMG1 VMG2 VMG5 VMSH1 VMSH3
			VMG2			
			VMG3			
			VMG4			
			VMG5			
			VMG6			
			VMG8			
			VMSI7			

4.2.8 MODELO LOGIT GLOBAL CON VARIABLES CUALITATIVAS SECTORIALES

En este apartado se aborda la hipótesis H_2 , es decir, si la introducción de variables cualitativas, representativas de los sectores económicos objeto de estudio, mejora el modelo global genérico estimado sin la consideración de dichas variables sectoriales.

Para dar respuesta a esta cuestión, se abordará, en primer lugar, la estimación de un modelo global elaborado con una muestra de todos los sectores económicos, para M.I., e introduciendo variables cualitativas representativas de los sectores. En segundo lugar, se compararan los resultados obtenidos con los correspondientes al modelo global estimado sin la participación de variables cualitativas.

En la tabla 4.33 se muestran los resultados del modelo Logit global estimado con variables cualitativas para M.I. En este modelo han resultado ser siete las variables seleccionadas: VMG1, VMG2, VMG3, VMG4, VMG6, la variable cualitativa o dummy del sector de la construcción y la dummy del sector de comercio y servicios.

Al analizar los signos de los coeficientes obtenidos para cada una de estas variables, se puede apreciar que las variables VMG1 y VMG2 presentan signo negativo, además del coeficiente de la dummy del sector de comercio y servicios. Por su parte, las variables VMG3,

VMG4, VMG6 y la dummy del sector de la construcción presentan signos positivos en sus coeficientes.

Si se analizan los Odds ratio de las variables, como cabía esperar, tan solo son mayores que 1 los correspondientes a las variables VMG3, VMG4, VMG6 y la dummy del sector de la construcción en tanto que son las únicas variables con signo positivo, siendo valores menores a 1 en el resto de los casos.

En el caso de VMG3 al tener signo positivo y un Odds ratio de 3,840 sugiere que cuando el ratio Fondo de Maniobra/Total Activo aumenta en un punto, la probabilidad de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 3,840 veces. Las variables con signo negativo (VMG1, VMG2 y la dummy de comercio y servicios) presentan un Odds ratio entre 0 y 1.

Por lo que se refiere a la significatividad de las variables incluidas en el modelo, todas las variables presentan niveles de significatividad por debajo del 5%, salvo la dummy del sector de la construcción, siendo, por tanto, relevantes en la explicación del modelo.

Una vez valorada la significación de cada variable de forma individual, procede analizar los distintos tests para constatar la robustez y precisión del modelo en su conjunto.

TABLA 4.33 MODELO LOGIT GLOBAL CON VARIABLES CUALITATIVAS. M.I

			Valor
Ajuste del modelo			
Significatividad Prueba Ómnibus			0,000
Significatividad Test de Hosmer y Lemeshow			0,400
-2 log verosimilitud (-2LL)			318,395
R ² Cox & Snell			0,458
R ² Nagelkerke			0,611
Matriz de clasificación			%
Dentro de la muestra			82,5
Insolventes			75,0
Solventes			89,9
Muestra de testeo			81,7
Insolventes			77,7
Solventes			85,6
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-28,574	0,000	0,000
VMG2	-1,189	0,305	0,001
VMG3	1,345	3,840	0,010
VMG4	13,108	492.718,43	0,003
VMG6	1,243	3,465	0,001
Dummy (Construcción)	61,492	5,080E+26	0,985
Dummy (Comercio y Servicios)	-0,665	0,514	0,021
Constante	-0,755	0,470	0,010

Notas: VMG1 = Beneficio después de Impuestos/Total Activo; VMG2 = Activo Corriente/Pasivo Corriente; VMG3 = Fondo de Maniobra/Total Activo; VMG4 = EBIT/Total Activo; VMG6 = Quick Ratio.

Por lo que se refiere a la prueba de Razón de Verosimilitud (RV) de todos los coeficientes, el valor

obtenido es de 0,000, concluyéndose que el ajuste es adecuado. En segundo lugar, el análisis de la capacidad predictiva del modelo mediante la matriz de clasificación nos indica que se han clasificado correctamente el 80,1% de las empresas, obteniéndose un mayor nivel de predicción en el caso de las empresas solventes, con un 86,2% de acierto, y clasificando correctamente el 73,7% de las empresas insolventes. Con respecto al nivel de clasificación en la muestra externa utilizada para comprobar la capacidad de generalización del modelo, este porcentaje ha resultado ser del 81,1%.

Además, la significatividad del estadístico de Hosmer y Lemeshow es 0,400 y, en consecuencia, se aceptaría la hipótesis de que no existen diferencias significativas entre las clasificaciones observadas y predichas. Por otro lado, el test $-2 \log$ de verosimilitud ($-2LL$) toma un valor de 318,395. Y finalmente, el coeficiente de Cox y Snell asciende a 0,331 y el R^2 de Nagelkerke alcanza un valor de 0,441, denotando ambos resultados también un ajuste aceptable del modelo estimado.

Una vez estimado y analizado el modelo global con variables cualitativas, procede compararlo con el modelo global obtenido sin la inclusión de variables de tipo cualitativo.

La comparación va a permitir dar respuesta a la segunda hipótesis de esta tesis que cuestionaba la superioridad de un modelo global sin dummies sectoriales

frente a un modelo que incluyera este tipo de variables representativa de los distintos sectores analizados.

Para llevar a cabo esta comparación se han utilizado tres criterios de selección de modelos, ya tratados en el capítulo 2. Estos criterios son los de Akaike, el bayesiano de Schwartz y el criterio de Hannann-Quinn. La regla de decisión para los tres criterios es la misma: el mejor modelo es el que ofrece el menor valor en la comparación.

TABLA 4.34 COMPARACIÓN DE MODELOS GLOBALES. M.I.

Ajuste del Modelo	Global SIN DUMMYS	Global CON DUMMYS
Prueba Ómnibus	0,000	0,000
Test Hosmer-Lemeshow	0,085	0,400
R ² Nagelkerke	0,441	0,611
Matriz de Clasificación (%)		
Fuera de la Muestra		
Solventes	86,9	85,6
Insolventes	74,5	77,7
Total	80,8	81,7
Criterio de Akaike	398,74	330,40
Criterio de Schwartz	405,34	341,35
Criterio de Hannan-Quinn	395,24	325,39

De acuerdo con esta regla, puede comprobarse a partir de la tabla 4.34 cómo el modelo global con dummies es superior al modelo global estimado sin la participación de variables cualitativas sectoriales.

De esta forma, queda aceptada también la hipótesis H_2 para los modelos estimados un año antes de la entrada en quiebra (M.I.). Es decir, la inclusión de variables cualitativas, representativas de los sectores económicos considerados, enriquece y aumenta la capacidad de explicación del modelo global.

4.2.9 ANÁLISIS DE ROBUSTEZ

En este apartado se llevará a cabo un análisis de la robustez de los resultados obtenidos en los diferentes modelos.

Este análisis constará de dos partes: por un lado, se ofrecerán los resultados de una validación cruzada de los modelos y se comprobará la capacidad del modelo global con variables cualitativas, en M.I., para conseguir buenos resultados utilizando las muestras de cada uno de los sectores. De otro, se aplicarán a los distintos modelos obtenidos los criterios de Akaike, de Schwartz y de Hanann-Quinn, todos ellos criterios informativos de selección de modelos.

Este análisis de robustez nos posibilitará contrastar la hipótesis H_3 , es decir, si el modelo global con dummies puede predecir correctamente con muestras de los diferentes sectores.

La validación cruzada (CV) es uno de los métodos más utilizados para evaluar la precisión predictiva de un modelo (Allen, 1974; Stone, 1974; Geisser, 1975).

Básicamente, consiste en dividir los datos muestrales, en los que una parte de los mismos se han empleado en la construcción de cada modelo, utilizando el resto de los datos para medir la precisión de los modelos estimados por medio de la evaluación de los errores de validación. Como resultado de este proceso, resultaría seleccionado el modelo con el mejor rendimiento global (Zhang y Yan, 2015).

CV se ha utilizado ampliamente en la minería de datos, bien sea para la selección de un modelo entre varios, o como un procedimiento para seleccionar la forma funcional del modelo (Hastie et al., 2009).

Las tablas 4.35 y 4.36 informan de la validación cruzada del modelo global con dummies en las distintas muestras sectoriales. Los resultados ponen de manifiesto, a tenor de lo que indican los diferentes criterios de selección, que el modelo global con dummies es capaz de predecir con éxito en las muestras sectoriales. Por tanto, la hipótesis H_3 queda aceptada.

TABLA 4.35 RESULTADOS DE LOS MODELOS DE REGRESIÓN LOGÍSTICA PARA M.I

Ajuste del Modelo	Global	Agricultura	Industria	Construcción	Comercio y Servicios	Hostelería
Prueba Ómnibus	0,000	0,000	0,000	0,000	0,000	0,000
Test Hosmer-Lemeshow	0,400	0,487	0,353	0,177	0,149	0,131
R ² Nagelkerke	0,611	0,560	0,657	0,559	0,584	0,690
Matriz de Clasificación (%)						
Fuera de la Muestra						
Solventes	85,6	57,9	95,1	78,6	79,5	90,9
Insolventes	77,7	90,9	82,0	84,1	86,5	91,5
Total	81,7	75,6	89,2	81,5	83,5	91,2
Criterio de Akaike	330,40	84,31	175,94	270,01	229,61	259,87
Criterio de Schwartz	341,67	87,96	190,97	278,69	240,41	268,98
Criterio de Hanann-Quinn	325,46	81,36	169,36	265,73	224,56	255,70

TABLA 4.36 ROBUSTEZ REGRESIÓN LOGÍSTICA DEL MODELO GLOBAL CON VARIABLES CUALITATIVAS PARA M.I.

Ajuste del Modelo	Agricultura	Industria	Construcción	Comercio y Servicios	Hostelería
Prueba Ómnibus	0,012	0,000	0,000	0,000	0,000
Test Hosmer-Lemeshow	0,562	0,589	0,757	0,884	0,975
R ² Nagelkerke	0,414	0,728	0,624	0,703	0,757
Matriz de Clasificación (%)					
Fuera de la Muestra					
Solventes	75,0	93,0	84,6	91,5	88,9
Insolventes	70,0	78,3	80,3	83,1	87,0
Total	72,7	85,8	82,5	87,3	88,0
Criterio de Akaike	56,34	112,69	128,25	117,10	98,44
Criterio de Schwartz	63,64	123,96	138,68	127,90	109,38
Criterio de Hanann-Quinn	50,43	107,75	123,12	112,05	93,43

4.3 RESULTADOS CON INFORMACIÓN DE DOS AÑOS ANTES DE LA QUIEBRA (M.II)

4.3.1 RESULTADOS MODELO GLOBAL

4.3.1.1 Análisis Exploratorio

Como se ha expuesto con anterioridad, en el análisis exploratorio se efectúa tanto un análisis descriptivo de los principales parámetros estadísticos de las variables, como la aplicación de tests paramétricos o no paramétricos y un análisis de correlación.

4.3.1.1.1 Análisis Descriptivo

Este análisis descriptivo se presenta de forma diferenciada para las empresas solventes e insolventes de la muestra global de empresas, en la que se encuentran incluidas empresas representativas de todos los sectores económicos considerados.

La tabla 4.37 detalla los resultados del análisis descriptivo realizado para la muestra global con información de dos años antes de la quiebra (M.II). Observando las medias obtenidas para cada una de las variables se comprueba que tanto VMG1 como VMG4 presentan medias con diferente signo según se trate de empresas solventes o insolventes.

Al observar las medias y las medianas de cada una de las variables, en la mayor parte de los casos, los

valores son más elevados en las empresas solventes que en el caso de las insolventes. Si se analizan los valores de las medias, el 56% de los valores medios de las empresas solventes superan al de las empresas insolventes. Para el caso de la mediana este porcentaje es del 67%.

En concreto, las variables VMG3, VMG7, VMG8 y VMSG9 presentan medias inferiores en las empresas solventes frente a las insolventes.

Por su parte, la desviación típica presenta valores moderados, salvo para VMG2, VMG5 y VMG6.

**TABLA 4.37 ESTADÍSTICOS DESCRIPTIVOS
(MODELO GLOBAL, M.II)**

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,05	-0,10	0,01	-0,00	0,21	0,50	-0,78	-9,24	5,26	2,02
	VMG2	3,39	2,74	1,29	1,05	11,80	20,69	0,00	0,00	242,72	579,92
	VMG3	0,24	0,37	0,22	0,35	0,30	0,38	-2,05	-2,74	0,97	0,99
	VMG4	0,05	-0,07	0,03	0,00	0,41	0,44	-6,77	-7,71	5,28	2,71
	VMG5	2,14	1,46	1,21	0,96	11,43	2,76	0,00	0,00	324,76	55,81
	VMG6	2,36	1,63	0,84	0,58	10,37	20,27	0,00	0,00	218,42	579,92
	VMG7	0,23	0,38	0,13	0,30	0,28	0,86	0,00	0,00	2,33	22,96
	VMG8	0,60	0,61	0,63	0,68	0,29	0,31	0,00	0,00	1,00	1,00
	VMG9	0,03	0,07	0,05	0,02	1,95	2,58	-41,71	-40,17	7,91	24,11

4.3.1.1.2 Análisis de Normalidad

A continuación procede analizar la normalidad de las distintas variables (tabla 4.38).

Como se puede comprobar, todas las variables en la muestra M.II., tanto para empresas solventes como insolventes, presentan un nivel de significación inferior al 5%. En consecuencia, se rechaza la hipótesis nula de normalidad.

**TABLA 4.38 ANÁLISIS DE NORMALIDAD
(MODELO GLOBAL, M.II)**

		M.II (2 AÑOS ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	0,327	0,000	0,307	0,000
	VMG2	0,382	0,000	0,447	0,000
	VMG3	0,071	0,000	0,068	0,000
	VMG4	0,346	0,000	0,287	0,000
	VMG5	0,425	0,000	0,309	0,000
	VMG6	0,405	0,000	0,467	0,000
	VMG7	0,211	0,000	0,408	0,000
	VMG8	0,093	0,000	0,120	0,000
	VMG9	0,392	0,000	0,293	0,000

NOTA: K-S: test de Kolmogorov-Smirnov

4.3.1.1.3 Tests no paramétricos

En la tabla 4.39 aparecen los resultados de la aplicación de los tests a la muestra correspondiente a dos años antes de la quiebra. Como puede observarse, según el test de Mann Whitney, se rechaza la hipótesis nula en 8 de las 9 variables consideradas, aceptándose tan solo la igualdad de distribución en la variable VMG8. Esta variable rechazó la hipótesis nula a un nivel de significación del 5%. No obstante, se rechaza la hipótesis nula de igualdad en las distribuciones mediante los resultados obtenidos por

la aplicación del test de Kolmogorov-Smirnov en todas las variables consideradas en la muestra global.

**TABLA 4.39 TESTS NO PARAMÉTRICOS
(MODELO GLOBAL, M.II)**

		M.II (2 AÑOS ANTES)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	-20,914	0,000	10,355	0,000
	VMG2	-7,450	0,000	3,657	0,000
	VMG3	8,108	0,000	4,231	0,000
	VMG4	-14,330	0,000	7,175	0,000
	VMG5	-5,121	0,000	2,296	0,000
	VMG6	-8,178	0,000	3,889	0,000
	VMG7	8,135	0,000	4,061	0,000
	VMG8	0,785	0,433	1,296	0,069
	VMG9	-2,739	0,006	5,011	0,000

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

4.3.1.1.4 Análisis de Correlación

Para finalizar el análisis exploratorio, se procederá al cálculo de las correlaciones entre las variables. A tal fin la tabla 4.40 muestra los coeficientes de correlación bilateral en la muestra global con datos de M.II. Los resultados indican que existe una elevada correlación entre las variables VMG2 y VMG6, con un coeficiente de correlación de Pearson de 0,972, si bien la elevada correlación de esta pareja de variables era previsible ante la similitud en el cálculo de las mismas.

**TABLA 4.40 ANÁLISIS DE CORRELACIÓN BILATERAL
(MODELO GLOBAL, M.II)**

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	INSOLV
VMG1	1	0,024	0,128	0,771	0,173	0,020	-0,026	0,037	-0,020	-0,203
VMG2	0,024	1	0,073	0,002	-0,019	0,972	0,020	0,073	-0,001	-0,027
VMG3	0,128	0,073	1	0,061	-0,067	0,028	0,029	0,556	-0,023	0,174
VMG4	0,771	0,002	0,061	1	0,157	-0,002	-0,022	0,007	-0,009	-0,145
VMG5	0,173	-0,019	-0,067	0,157	1	-0,011	-0,022	0,057	-0,044	-0,036
VMG6	0,020	0,972	0,028	-0,002	-0,011	1	0,066	0,048	0,001	-0,026
VMG7	-0,026	0,020	0,029	-0,022	-0,022	0,066	1	-0,004	0,004	0,077
VMG8	0,037	0,073	0,556	0,007	0,057	0,048	-0,004	1	-0,011	0,011
VMG9	-0,020	-0,001	-0,023	-0,009	-0,044	0,001	0,004	-0,011	1	0,000
INSOLV	-0,203	-0,027	0,174	-0,145	-0,036	-0,026	0,077	0,011	0,000	1

4.3.1.2 Análisis Confirmatorio

En este apartado, con la muestra global de empresas, se estima un modelo de predicción utilizando información de dos años antes de producirse la insolvencia.

Asimismo, el modelo estimado será validado en una muestra externa. Los elementos de la muestra que se utilizarán en cada caso para la estimación de los modelos será de un 70% de las empresas para la denominada muestra de entrenamiento y el 30% restante para la denominada muestra externa o de testeo.

4.3.1.2.1 Modelo Logit Global sin variables cualitativas

En la tabla 4.41 se muestran los resultados del modelo Logit global sin dummies estimado para M.II. Las variables finalmente seleccionadas han resultado ser cuatro: VMG1, VMG2, VMG3 y VMG9.

Al analizar los signos de los coeficientes obtenidos para cada una de estas variables se puede apreciar que las variables VMG1, VMG2 y VMG9 presentan signo negativo, lo cual denota una relación inversa con la insolvencia. Por su parte, la variable VMG3 presenta signo positivo en su coeficiente. En este caso, la relación con la insolvencia es directa.

Considerando, además, los Odds ratio de las variables, tan solo son mayores que 1 en el correspondiente a la variable VMG3, en tanto que es la única variable con signo positivo, siendo menores a 1 en el resto de los casos.

Es destacable la importancia de la variable VMG3, pues con un Odds ratio de 4,269 indica que cuando el ratio Fondo de Maniobra/Total Activo aumenta en un punto, la probabilidad de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 4,269 veces.

Por lo que se refiere a la significatividad de las variables incluidas en el modelo, todas las variables

presentan niveles de significatividad por debajo del 5%, siendo, por tanto, relevantes en la explicación del modelo.

TABLA 4.41 MODELO LOGIT GLOBAL SIN VARIABLES CUALITATIVAS. M.II

Ajuste del modelo		Valor	
Significatividad Prueba Ómnibus		0,000	
Significatividad Test de Hosmer y Lemeshow		0,139	
-2 log verosimilitud (-2LL)		453,922	
R ² Cox & Snell		0,269	
R ² Nagelkerke		0,359	
Matriz de clasificación		%	
Dentro de la muestra		73,0	
Insolventes		72,2	
Solventes		73,9	
Muestra de testeo		71,1	
Insolventes		69,9	
Solventes		72,3	
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-14,456	0,000	0,000
VMG2	-0,080	0,924	0,030
VMG3	1,451	4,269	0,000
VMG9	-0,400	0,670	0,016
Constante	-0,159	0,853	0,369

Notas: VMG1 = Beneficio después de Impuestos/Total Activo; VMG2 = Activo Corriente/Pasivo Corriente; VMG3 = Fondo de Maniobra/Total Activo; VMG9 = Beneficio después de Impuestos/Patrimonio Neto.

Una vez valorada la significación de cada variable de forma individual, procede comprobar la robustez y la precisión del modelo en su conjunto. En tal sentido, la prueba Ómnibus arroja un valor de 0,000. La matriz de

clasificación indica que se han clasificado correctamente el 73,0% de las empresas (71,1% para la muestra de testeo). Además, el test de Hosmer y Lemeshow arroja una significación de 0,139 y el R^2 de Nagelkerke es 0,359. Los resultados, pues, denotan un ajuste aceptable del modelo seleccionado.

4.3.2 RESULTADOS MODELO AGRICULTURA

4.3.2.1 Análisis Exploratorio

En este apartado se aborda el análisis exploratorio de las variables correspondientes a la muestra del sector agricultura.

4.3.2.1.1 Análisis Descriptivo

Los resultados que aparecen en la tabla 4.42 indican que los valores medios de todas las variables muestran un mismo signo positivo. Además, que las variables VMG3, VMG7, VMG8, VMSG9 Y VMSA3 presentan medias inferiores en el caso de las empresas solventes, y que la desviación típica de todas las variables es relativamente moderada.

**TABLA 4.42 ESTADÍSTICOS DESCRIPTIVOS
(MODELO AGRICULTURA, M.II)**

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,23	0,01	0,01	0,00	0,02	0,03	-0,03	-0,05	0,08	0,07
	VMG2	1,81	1,64	1,34	1,30	1,34	1,10	0,02	0,22	5,95	4,66
	VMG3	0,20	0,28	0,14	0,23	0,27	0,28	-0,62	-0,20	0,69	0,89
	VMG4	0,04	0,03	0,03	0,02	0,03	0,04	-0,00	-0,05	0,12	0,13
	VMG5	0,89	0,64	0,76	0,33	0,66	0,75	0,09	0,05	2,77	3,08
	VMG6	1,28	0,95	0,84	0,55	1,34	0,97	0,02	0,00	5,86	4,39
	VMG7	0,35	0,41	0,37	0,40	0,20	0,21	0,00	0,08	0,73	0,87
	VMG8	0,44	0,47	0,35	0,44	0,28	0,32	0,01	0,04	0,98	1,00
	VMG9	0,08	0,66	0,04	0,02	0,07	0,24	0,01	-0,57	0,26	0,79
VARIABLES MODELO SECTORIAL (AGRICULTURA)	VMSA1	2,52	0,96	1,00	0,47	4,31	1,19	-0,32	0,06	22,98	4,40
	VMSA2	2,85	1,88	1,76	1,21	2,24	2,93	-0,28	-5,80	8,12	10,81
	VMSA3	0,08	0,09	0,04	0,06	0,17	0,10	-0,01	-0,09	1,00	0,36

4.3.2.1.2 Análisis de Normalidad

En la tabla 4.43 se muestran los resultados del test de Kolmogorov-Smirnov para las empresas solventes e insolventes pertenecientes al sector agricultura, a un nivel de significación del 5%. Con la excepción de VMG3 para el caso de las empresas insolventes, todas las variables presentan un nivel de significación inferior al 5%. Por tanto, se rechaza la hipótesis nula de normalidad.

**TABLA 4.43 ANÁLISIS DE NORMALIDAD
(SECTOR AGRICULTURA, M.II)**

		M.II (2 AÑOS ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	0,203	0,000	0,301	0,000
	VMG2	0,271	0,000	0,233	0,000
	VMG3	0,130	0,012	0,095	0,200
	VMG4	0,152	0,001	0,227	0,000
	VMG5	0,151	0,001	0,201	0,000
	VMG6	0,243	0,000	0,240	0,000
	VMG7	0,155	0,001	0,114	0,047
	VMG8	0,121	0,026	0,110	0,057
	VMG9	0,192	0,000	0,285	0,000
VARIABLES MODELO SECTORIAL (AGRICULTURA)	VMSA1	0,246	0,000	0,258	0,000
	VMSA2	0,344	0,000	0,189	0,000
	VMSA3	0,421	0,000	0,340	0,000

NOTA: K-S: test de Kolmogorov-Smirnov

4.3.2.1.3 Tests no paramétricos

Según los resultados obtenidos tras la aplicación de tests no paramétricos a la muestra correspondiente a dos años antes de la quiebra (tabla 4.44), se acepta la hipótesis nula de igualdad de distribuciones para VMG3, VMG5, VMG7, VMG8 y VMG9.

**TABLA 4.44 TESTS NO PARAMÉTRICOS
(SECTOR AGRICULTURA, M.II)**

		M.II (2 AÑOS ANTES)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	-4,070	0,000	2,532	0,000
	VMG2	-3,287	0,001	1,875	0,002
	VMG3	2,433	0,015	1,305	0,066
	VMG4	-3,616	0,000	2,037	0,000
	VMG5	-0,919	0,358	0,815	0,520
	VMG6	-2,798	0,005	1,604	0,012
	VMG7	2,346	0,019	1,268	0,080
	VMG8	0,425	0,671	0,575	0,895
	VMG9	-0,019	0,985	2,286	0,000
VARIABLES MODELO SECTORIAL (AGRICULTURA)	VMSA1	-4,581	0,000	2,228	0,000
	VMSA2	-3,585	0,000	1,960	0,001
	VMSA3	-3,603	0,000	2,117	0,000

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

4.3.2.1.4 Análisis de Correlación

La tabla 4.45 muestra el análisis de correlación bilateral para la muestra correspondiente al sector agricultura, con información de dos años antes de la quiebra. Puede comprobarse que existe una elevada correlación entre las variables VMG1 y VMG4, en concreto, con un coeficiente de correlación de Pearson de 0,946. Asimismo, entre las parejas VMG2 y VMG6 (0,900) y entre VMSA2 y VMSA3 (0,885).

**TABLA 4.45 ANÁLISIS DE CORRELACIÓN BILATERAL
(SECTOR AGRICULTURA, M.II)**

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	VMSA1	VMSA2	VMSA3	INSOLV
VMG1	1	0,143	0,133	0,946	0,147	0,147	-0,148	0,022	-0,066	0,255	0,299	0,346	-0,262
VMG2	0,143	1	-0,030	0,114	-0,171	0,900	-0,197	0,068	-0,094	0,258	0,134	0,184	-0,293
VMG3	0,133	-0,030	1	-0,060	0,231	-0,130	0,019	0,492	0,114	-0,110	-0,089	0,035	0,196
VMG4	0,946	0,114	-0,060	1	0,076	0,124	-0,105	0,100	0,041	0,254	0,394	0,411	-0,236
VMG5	0,147	-0,171	0,231	0,076	1	-0,170	-0,114	0,464	0,181	0,006	-0,040	0,110	-0,009
VMG6	0,147	0,900	-0,130	0,124	-0,170	1	-0,200	0,002	-0,072	0,230	0,183	0,187	-0,273
VMG7	-0,148	-0,197	0,019	-0,105	-0,114	-0,200	1	-0,143	-0,057	-0,411	-0,223	0,018	0,224
VMG8	0,022	0,068	0,483	0,100	0,464	0,002	-0,143	1	0,147	-0,079	-0,116	0,166	0,037
VMG9	-0,066	-0,094	0,114	0,041	0,181	-0,072	-0,057	0,147	1	-0,188	-0,018	-0,024	0,156
VMSA1	0,255	0,258	-0,110	0,254	0,006	0,230	-0,411	-0,079	-0,188	1	0,385	0,137	-0,396
VMSA2	0,299	0,134	-0,089	0,394	-0,040	0,183	-0,223	-0,116	-0,018	0,385	1	0,207	-0,263
VMSA3	0,346	0,184	0,035	0,411	0,110	0,187	0,018	0,166	-0,024	0,137	0,207	1	-0,210
INSOLV	-0,262	-0,293	0,196*	-0,236	-0,009	-0,273	0,224	0,037	0,156	-0,396	-0,263	-0,210	1

4.3.2.2 Análisis Confirmatorio

En este apartado se realizará un análisis confirmatorio de los resultados obtenidos para el sector agricultura utilizando información de dos años previos a la situación de insolvencia (M.II).

4.3.2.2.1 Modelo Logit Sector Agricultura

Los resultados del Logit estimado para la muestra de empresas del sector agricultura con M.II aparecen en la tabla 4.46. Las variables estadísticamente significativas han sido VMG1 y VMSA1, ambas con signo negativo en la regresión.

La prueba Ómnibus ofrece un valor de 0,000. A su vez, la matriz de clasificación indica que se han clasificado correctamente el 65,2% de las empresas (69,6% en la muestra externa). Por su parte, el estadístico del test de Hosmer y Lemeshow es 0,210 y, el R^2 de Nagelkerke alcanza 0,327.

Los resultados, por tanto, denotan un ajuste aceptable del modelo construido.

**TABLA 4.46 MODELO LOGIT
SECTOR AGRICULTURA. M.II**

Ajuste del modelo	Valor		
Significatividad prueba Ómnibus	0,000		
Significatividad Test de Hosmer y Lemeshow	0,210		
-2 log verosimilitud (-2LL)	98,333		
R ² Cox & Snell	0,245		
R ² Nagelkerke	0,327		
Matriz de clasificación	%		
Dentro de la muestra	65,2		
	Insolventes	73,3	
	Solventes	56,8	
Muestra de testeo	69,6		
	Insolventes	88,5	
	Solventes	45,0	
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-13,429	0,000	0,045
VMSA1	-0,417	0,659	0,009
Constante	0,804	2,233	0,013

NOTA: VMG1 = Beneficio después de Impuestos/Total Activo; VMSA1 = Fondos Propios/Total Deuda

4.3.3 RESULTADOS MODELO INDUSTRIA

4.3.3.1 Análisis Exploratorio

En este apartado se aborda el análisis exploratorio para la muestra de empresas pertenecientes al sector industria.

4.3.3.1.1 Análisis Descriptivo

En la tabla 4.47 se detalla el análisis descriptivo realizado para la muestra M.II del sector industrial. A tal efecto se comprueba que las variables VMG3, VMG5, VMG8, VMSI3 y VMSI8 presentan valores medios inferiores en el colectivo de empresas solventes, en comparación a las insolventes. Así mismo, la desviación típica presenta valores moderados en todas las variables, salvo para VMSI2 y VMSI4.

**TABLA 4.47 ESTADÍSTICOS DESCRIPTIVOS
(MODELO INDUSTRIA, M.II)**

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,03	-0,03	0,01	0,00	0,03	0,08	-0,03	-0,24	0,14	0,07
	VMG2	1,72	1,24	1,34	1,04	1,29	1,02	0,18	0,16	6,95	7,65
	VMG3	0,27	0,51	0,25	0,48	0,21	0,23	-0,31	0,04	0,83	0,95
	VMG4	0,05	-0,00	0,03	0,02	0,04	0,09	-0,04	-0,21	0,20	0,17
	VMG5	1,30	1,32	1,16	1,18	0,71	0,80	0,15	0,11	4,05	3,87
	VMG6	1,11	0,76	0,80	0,61	1,10	0,70	0,05	0,03	5,42	4,76
	VMG7	0,28	0,23	0,26	0,21	0,19	0,21	0,00	0,00	0,87	0,80
	VMG8	0,54	0,60	0,54	0,62	0,22	0,21	0,08	0,16	0,97	0,96
	VMG9	0,09	-0,92	0,05	0,02	0,18	5,39	-1,12	-40,17	0,80	2,62
VARIABLES MODELO SECTORIAL (INDUSTRIA)	VMSI1	0,04	-0,03	0,03	0,01	0,04	0,15	-0,05	-0,84	0,26	0,12
	VMSI2	11,80	7,41	5,66	4,72	19,41	7,37	1,47	0,23	146,59	34,65
	VMSI3	-0,27	0,03	0,38	0,46	3,70	2,61	-35,17	-9,29	9,75	4,74
	VMSI4	57,63	5,64	2,25	0,63	239,0	22,00	-0,96	-14,90	1837,7	140,74
	VMSI5	0,01	0,03	0,00	0,02	0,02	0,02	0,00	0,00	0,19	0,14
	VMSI6	7,40	6,75	7,20	6,78	1,73	0,95	4,26	4,19	12,76	8,29
	VMSI7	0,14	-0,49	0,12	0,15	0,61	5,53	-6,02	-41,02	2,21	4,09
	VMSI8	4,82	5,85	2,74	4,32	5,74	6,23	0,20	0,17	34,98	30,23

4.3.3.1.2 Análisis de Normalidad

Los resultados sobre normalidad indican que, con la excepción de VMG3, todas las variables en M.II tienen un nivel de significación inferior al 5%. Esto implica rechazar la hipótesis nula de normalidad (tabla 4.48).

**TABLA 4.48 ANÁLISIS DE NORMALIDAD
(MODELO INDUSTRIA, M.II)**

		M.II (2 AÑOS ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	0,286	0,000	0,206	0,000
	VMG2	0,325	0,000	0,467	0,000
	VMG3	0,050	0,200	0,089	0,000
	VMG4	0,232	0,000	0,191	0,000
	VMG5	0,188	0,000	0,299	0,000
	VMG6	0,295	0,000	0,465	0,000
	VMG7	0,155	0,000	0,447	0,000
	VMG8	0,066	0,026	0,136	0,000
	VMG9	0,453	0,000	0,340	0,000
VARIABLES MODELO SECTORIAL (INDUSTRIA)	VMSI1	0,186	0,000	0,506	0,000
	VMSI2	0,371	0,000	0,386	0,000
	VMSI3	0,291	0,000	0,338	0,000
	VMSI4	0,413	0,000	0,446	0,000
	VMSI5	0,271	0,000	0,324	0,000
	VMSI6	0,072	0,010	0,035	0,000
	VMSI7	0,427	0,000	0,329	0,000
	VMSI8	0,336	0,000	0,404	0,000

4.3.3.1.3 Tests no paramétricos

Los tests aplicados a la muestra M.II indican que tan sólo VMG5, VMSI3 y VMSI7 registran un nivel de

significación superior al 5% ($p > 0,05$), con lo cual se aceptaría solo en estos casos la hipótesis nula de igualdad de distribuciones (tabla 4.49).

**TABLA 4.49 TESTS NO PARAMÉTRICOS
(MODELO INDUSTRIA, M.II)**

		M.II (2 AÑOS ANTES)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	-8,425	0,000	4,801	0,000
	VMG2	-4,475	0,000	2,406	0,000
	VMG3	5,988	0,000	3,307	0,000
	VMG4	-7,286	0,000	4,694	0,000
	VMG5	-1,180	0,238	1,761	0,004
	VMG6	-2,877	0,004	1,541	0,017
	VMG7	-4,479	0,000	3,086	0,000
	VMG8	3,163	0,002	2,245	0,000
	VMG9	-2,049	0,040	3,390	0,000
VARIABLES MODELO SECTORIAL (INDUSTRIA)	VMSI1	-7,137	0,000	4,420	0,000
	VMSI2	-4,648	0,000	2,921	0,000
	VMSI3	0,579	0,563	1,350	0,052
	VMSI4	-6,394	0,000	2,822	0,000
	VMSI5	3,487	0,000	2,423	0,000
	VMSI6	-3,066	0,003	1,854	0,002
	VMSI7	-0,712	0,476	2,837	0,000
	VMSI8	1,647	0,100	1,643	0,009

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

4.3.3.1.4 Análisis de Correlación

Para finalizar el análisis exploratorio del sector industrial en M.II se procederá al cálculo de las correlaciones que se producen entre las variables.

En la tabla 4.50 se muestra el análisis de correlación bilateral conforme al coeficiente de Pearson. En este caso se detecta la existencia de una elevada correlación entre los pares de variables VMG2 y VMG6 (0,998); VMG9 y VMSI7 (0,964); VMG1 y VMG4 (0,953); VMG7 y VMSI1 (-0,979); y VMSI8 y VMG5 (0,785).

**TABLA 4.50 ANÁLISIS DE CORRELACIÓN BILATERAL
(MODELO INDUSTRIA, M.II)**

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	VMSI1	VMSI2	VMSI3	VMSI4	VMSI5	VMSI6	VMSI7	VMSI8	INSOLV
VMG1	1	0,004	0,149	0,953	0,112	-0,002	-0,073	0,029	-0,193	0,097	0,303	-0,027	0,024	-0,127	0,113	-0,180	0,047	-0,322
VMG2	0,004	1	0,088	0,001	-0,068	0,998	0,020	0,078	0,002	-0,237	-0,021	0,011	0,003	-0,091	0,029	-0,003	-0,011	0,033
VMG3	0,149	0,088	1	0,144	0,066	0,083	0,011	0,596	-0,027	0,000	-0,207	0,021	-0,089	-0,016	-0,028	-0,014	0,163	0,229
VMG4	0,953	0,001	0,144	1	0,134	-0,006	-0,061	0,030	-0,121	0,064	0,105	-0,034	0,002	-0,080	0,134	-0,144	0,069	-0,310
VMG5	0,112	-0,068	0,066	0,134	1	-0,069	-0,057	0,132	0,001	0,034	0,167	0,028	0,021	-0,123	-0,335	0,007	0,785	0,019
VMG6	-0,002	0,998	0,083	-0,006	-0,069	1	0,082	0,078	0,002	-0,328	-0,022	0,010	0,000	-0,103	0,033	-0,003	-0,014	0,043
VMG7	-0,073	0,020	0,011	-0,061	-0,057	0,082	1	0,063	0,007	-0,979	-0,027	0,004	-0,034	0,052	0,111	0,002	-0,134	0,056
VMG8	0,029	0,078	0,596	0,030	0,132	0,078	0,063	1	-0,078	-0,069	-0,123	0,085	-0,035	-0,162	-0,155	-0,080	0,251	0,12
VMG9	-0,193	0,002	-0,027	-0,121	0,001	0,002	0,007	-0,078	1	-0,002	-0,322	0,195	-0,001	0,028	0,020	0,964	0,014	-0,017
VMSI1	0,097	-0,237	0,000	0,064	0,034	-0,328	-0,979	-0,069	-0,002	1	0,016	0,000	0,014	0,034	-0,089	0,001	0,004	-0,054
VMSI2	0,303	-0,021	-0,207	0,105	0,167	-0,022	-0,027	-0,123	-0,322	0,016	1	0,032	0,103	-0,060	-0,218	-0,258	0,038	-0,068
VMSI3	-0,027	0,011	0,021	-0,034	0,028	0,010	0,004	0,085	0,195	0,000	0,032	1	0,000	0,018	0,026	0,282	0,020	-0,036
VMSI4	0,024	0,003	-0,089	0,002	0,021	0,000	-0,034	-0,035	-0,001	0,014	0,103	0,000	1	-0,071	0,026	-0,008	-0,023	-0,101
VMSI5	-0,127	-0,091	-0,016	-0,080	-0,123	-0,103	0,052	-0,162	0,028	0,034	-0,060	0,018	-0,071	1	0,052	0,051	-0,045	0,128
VMSI6	0,113	0,029	-0,028	0,134	-0,335	0,033	0,111	-0,155	0,020	-0,089	-0,218	0,026	0,026	0,052	1	0,010	-0,199	-0,179
VMSI7	-0,180	-0,003	-0,014	-0,144	0,007	-0,003	0,002	-0,080	0,964	0,001	-0,258	0,282	-0,008	0,051	0,010	1	0,009	0,027
VMSI8	0,047	-0,011	0,163	0,069	0,785	-0,014	-0,134	0,251	0,014	0,004	0,038	0,020	-0,023	-0,045	-0,199	0,009	1	-0,121
INSOLV	-0,322	0,033	0,229	-0,310	0,019	0,043	0,056	0,122	-0,017	-0,054	-0,068	-0,036	-0,101	0,128	-0,179	0,027	0,121	1

4.3.3.2 Análisis Confirmatorio

4.3.3.2.1 Modelo Logit Sector Industria

Las variables seleccionadas por el modelo Logit estimado han sido VMG1, VMG2, VMG3, VMG4, VMG6, VMG7, VMG8, VMG9 y VMSI4. De ellas, VMG1, VMG2, VMG7, VMG8, VMG9 y VMSI4 presentan signo negativo en el coeficiente de regresión, mientras que las variables VMG3, VMG4 y VMG6 lo presentan positivo. Además, destaca la variable VMG6 con un Odds Ratio de 36,692, indicando que cuando aumenta en un punto, la probabilidad de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 36,692 veces (tabla 4.51).

Respecto al modelo en su conjunto, los resultados denotan un ajuste aceptable, pues tanto la prueba Ómnibus como el test de Hosmer y Lemeshow y el R^2 de Nagelkerke (0,832) alcanzan valores adecuados.

Además, el modelo ha clasificado correctamente el 94,5% de las empresas, obteniéndose un mayor nivel de predicción en el caso de las empresas solventes, con un 95,8% de acierto, y clasificando correctamente el 91,4% de las empresas insolventes. Por su parte, el nivel de clasificación en la muestra externa, utilizada para comprobar la capacidad de generalización del modelo, ha resultado ser del 78,1%.

TABLA 4.51 MODELO LOGIT
SECTOR INDUSTRIAL. M.II

Ajuste del modelo		Valor	
Significatividad prueba Ómnibus		0,000	
Significatividad Test de Hosmer y Lemeshow		0,189	
-2 log verosimilitud (-2LL)		66,261	
R ² Cox & Snell		0,582	
R ² Nagelkerke		0,832	
Matriz de clasificación		%	
Dentro de la muestra		94,5	
Insolventes		91,4	
Solventes		95,8	
Muestra de testeo		78,1	
Insolventes		72,5	
Solventes		83,3	
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-105,627	0,000	0,002
VMG2	-2,581	0,076	0,017
VMG3	21,487	2.146.707.735	0,000
VMG4	61,678	6,115E+26	0,003
VMG6	3,603	36,692	0,012
VMG7	-9,142	0,000	0,002
VMG8	-12,287	0,000	0,004
VMG9	-3,351	0,035	0,031
VMSI4	-0,039	0,962	0,041
Constante	-1,413	0,243	0,614

NOTA: VMG1 = Beneficio después de Impuestos/Total Activo; VMG2 = Activo Corriente/Pasivo Corriente; VMG3 = Fondo de Maniobra/Total Activo; VMG4 = EBIT/Total Activo; VMG6 = Quick Ratio; VMG7 = Total Deuda/Total Activo; VMG8 = Activo Corriente/Total Activo; VMG9 = Beneficio después de Impuestos/Patrimonio Neto; VMSI4 = Fondos Propios/Pasivo No Corriente.

4.3.4 RESULTADOS MODELO CONSTRUCCIÓN

4.3.4.1 Análisis Exploratorio

4.3.4.1.1 Análisis Descriptivo

Para el sector de construcción, los principales estadísticos de las variables aparecen en la tabla 4.52. A tal efecto se comprueba que tan sólo VMG1 y VMG9 presentan medias con diferente signo, según se trate de empresas solventes e insolventes.

**TABLA 4.52 ESTADÍSTICOS DESCRIPTIVOS
(MODELO CONSTRUCCIÓN, M.II)**

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,04	-0,01	0,01	0,00	0,07	0,07	-0,04	-0,34	0,39	0,32
	VMG2	4,28	3,52	1,71	1,62	8,12	5,86	0,00	0,13	55,70	52,96
	VMG3	0,34	0,62	0,29	0,71	0,30	0,29	-0,39	-0,23	0,97	0,99
	VMG4	-0,01	0,01	0,00	0,02	0,43	0,08	-2,14	-0,33	2,02	0,43
	VMG5	0,63	0,79	0,32	0,49	0,74	0,80	0,00	0,00	3,23	3,98
	VMG6	2,49	0,77	0,92	0,62	5,12	0,71	0,00	0,00	30,68	5,01
	VMG7	0,34	0,48	0,29	0,50	0,29	0,29	0,00	0,00	1,92	1,33
	VMG8	0,60	0,82	0,67	0,89	0,31	0,20	0,00	0,04	1,00	1,00
	VMG9	0,19	0,08	0,04	0,03	0,90	0,96	-4,45	-4,27	7,63	7,39
VARIABLES MODELO SECTORIAL (CONSTRUCCIÓN)	VMSC1	0,09	0,60	0,00	0,54	1,13	1,40	-4,65	-5,56	9,90	9,48
	VMSC2	0,38	-0,15	0,09	0,03	0,99	2,50	-2,12	-20,20	6,96	9,72
	VMSC3	4,99	1,11	1,07	0,23	10,53	6,02	-0,77	-5,87	64,10	77,15

Además, se puede apreciar cómo el 67% de los valores medios en las empresas solventes supera al de las

empresas insolventes. Para el caso de la mediana, este porcentaje es del 44%. En concreto, las variables VMG3, VMG5, VMG7, VMG8 y VMSC2 presentan medias inferiores en el caso de las empresas solventes frente a las insolventes. Por su parte, la desviación típica presenta valores aceptables, salvo para VMG2 y VMG6.

4.3.4.1.2 Análisis de Normalidad

Todas las variables presentan un nivel de significación inferior al 5% en el test de Kolmogorov-Smirnov, lo que implica rechazar la hipótesis nula de normalidad (tabla 4.53).

**TABLA 4.53 ANÁLISIS DE NORMALIDAD
(MODELO CONSTRUCCIÓN, M.II)**

		M.II (2 AÑOS ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	0,301	0,000	0,180	0,000
	VMG2	0,450	0,000	0,279	0,000
	VMG3	0,076	0,005	0,134	0,000
	VMG4	0,335	0,000	0,220	0,000
	VMG5	0,326	0,000	0,201	0,000
	VMG6	0,461	0,000	0,292	0,000
	VMG7	0,213	0,000	0,085	0,001
	VMG8	0,148	0,000	0,207	0,000
	VMG9	0,332	0,000	0,246	0,000
VARIABLES MODELO SECTORIAL (CONSTRUCCIÓN)	VMSC1	0,265	0,000	0,230	0,000
	VMSC2	0,358	0,000	0,451	0,000
	VMSC3	0,309	0,000	0,364	0,000

NOTA: K-S: test de Kolmogorov-Smirnov

4.3.4.1.3 Tests no paramétricos

A continuación se aplicará a cada una de las variables distintos tests no paramétricos. Se aplicaran dos tests. Por un lado el test de Mann-Whitney y, por otro, el test de Kolmogorov-Smirnov.

Las variables VMG3 y VMG5 muestran un nivel de significación superior al 5% ($p > 0,05$) en los tests de Mann-Whitney y Kolmogorov-Smirnov, por lo que se aceptaría sólo en estos casos la hipótesis nula de igualdad de distribuciones (tabla 4.54).

**TABLA 4.54 TESTS NO PARAMÉTRICOS
(MODELO CONSTRUCCIÓN, M.II)**

		M.II (2 AÑOS ANTES)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	-10,057	0,000	4,286	0,000
	VMG2	0,218	0,827	0,667	0,765
	VMG3	8,847	0,000	4,294	0,000
	VMG4	1,077	0,281	2,346	0,000
	VMG5	-1,423	0,155	0,847	0,470
	VMG6	-5,511	0,000	3,037	0,000
	VMG7	7,298	0,000	3,540	0,000
	VMG8	4,752	0,000	2,586	0,000
	VMG9	-1,823	0,068	2,255	0,000
VARIABLES MODELO SECTORIAL (CONSTRUCCIÓN)	VMSC1	5,186	0,000	3,642	0,000
	VMSC2	-6,432	0,000	2,742	0,000
	VMSC3	-6,862	0,000	3,525	0,000

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

4.3.4.1.4 Análisis de Correlación

Para finalizar el análisis exploratorio del sector construcción en M.II se procederá al cálculo de las correlaciones que se producen entre las variables, considerando cada pareja de variables posible, incluida, propiamente, la variable insolvencia. Conforme a los resultados que aparecen en la 4.55, el nivel de correlación existente no es elevado.

**TABLA 4.55 ANÁLISIS DE CORRELACIÓN BILATERAL
(SECTOR CONSTRUCCIÓN, M.II)**

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	VMSC1	VMSC2	VMSC3	INSOLV
VMG1	1	0,033	-0,072	-0,052	0,627	0,040	-0,280	0,080	-0,100	-0,004	0,069	0,210	-0,292
VMG2	0,033	1	-0,024	0,006	-0,031	0,996	-0,026	0,043	-0,015	-0,004	0,008	0,115	-0,060
VMG3	-0,072	-0,024	1	-0,001	-0,133	-0,040	0,305	0,582	-0,070	0,160	-0,044	-0,267	0,427
VMG4	-0,052	0,006	-0,001	1	-0,014	0,006	0,051	-0,055	-0,017	0,052	0,059	-0,072	0,049
VMG5	0,627	-0,031	-0,133	-0,014	1	-0,016	-0,275	0,048	-0,049	-0,029	0,016	-0,016	-0,131
VMG6	0,040	0,996	-0,040	0,006	-0,016	1	-0,055	0,041	0,002	-0,011	0,007	0,339	-0,062
VMG7	-0,280	-0,026	0,305	0,051	-0,275	-0,055	1	-0,026	-0,125	0,082	-0,051	-0,362	0,300
VMG8	0,080	0,043	0,582	-0,055	0,048	0,041	-0,026	1	-0,024	0,135	-0,025	-0,107	0,265
VMG9	-0,100	-0,015	-0,070	-0,017	-0,049	0,002	-0,125	-0,024	1	0,008	-0,026	0,026	-0,018
VMSC1	-0,004	-0,004	0,160	0,052	-0,029	-0,011	0,082	0,135	0,008	1	0,019	-0,090	0,156
VMSC2	0,069	0,008	-0,044	0,059	0,016	0,007	-0,051	-0,025	-0,026	0,019	1	0,053	-0,094
VMSC3	0,210	0,115	-0,267	-0,072	-0,016	0,339	-0,362	-0,107	0,026	-0,090	0,053	1	-0,253
INSOLV	-0,292	-0,060	0,427	0,049	-0,131	-0,062	0,300	0,265	-0,018	0,156	-0,094	-0,253	1

4.3.4.2 Análisis Confirmatorio

Como en casos anteriores, para la construcción del modelo Logit se ha utilizado el 70% de los datos de la muestra. Posteriormente, el modelo se ha testado con el 30% de los restantes datos muestrales.

4.3.4.2.1 Modelo Logit Sector Construcción

En el modelo estimado para el sector construcción, utilizando información de dos años antes de la quiebra (M.II), las variables seleccionadas son VMG1, VMG3, VMG6, VMG7, VMG8 y VMSC1 (tabla 4.56).

Además, las variables VMG1 y VMG6 presentan signo negativo, lo cual denota una relación inversa con la insolvencia. En este caso, el Odds Ratio vale 0,000 para VMG1, lo cual indica que, si bien con respecto a la insolvencia no es una variable relevante, sí lo es con respecto a las empresas solventes.

Por su parte, VMG7 tiene un Odds ratio de 3,479, lo que significa que cuando el ratio Total Deuda/Total Activo aumenta en un punto, la probabilidad de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 3,479 veces.

Con respecto a la significatividad de los coeficientes del modelo, puede apreciarse que todos ellos son significativos con un p-valor inferior o igual al 5%.

TABLA 4.56 MODELO LOGIT
SECTOR CONSTRUCCIÓN. M.II

Ajuste del modelo		Valor	
Prueba Ómnibus para todos los coeficientes		0,000	
Significatividad Test de Hosmer y Lemeshow		0,117	
-2 log verosimilitud (-2LL)		361,278	
R ² Cox & Snell		0,372	
R ² Nagelkerke		0,496	
Matriz de clasificación		%	
Dentro de la muestra		77,0	
Insolventes		78,7	
Solventes		75,4	
Muestra de testeo		74,5	
Insolventes		72,1	
Solventes		76,6	
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-16,729	0,000	0,000
VMG3	1,775	5,902	0,001
VMG6	-0,139	0,870	0,022
VMG7	1,247	3,479	0,024
VMG8	1,584	4,874	0,014
VMSC1	0,289	1,336	0,014
Constante	-2,006	0,135	0,000

Notas: VMG1 = Beneficio después de Impuestos/Total Activo; VMG3 = Fondo de Maniobra/Total Activo; VMG6 = Quick Ratio; VMG7 = Total Deuda/Total Activo; VMG8 = Activo Corriente/Total Activo; VMSC1 = Gastos Financieros/EBIT.

Respecto a la robustez y a la precisión del modelo, la prueba de Razón de Verosimilitud (RV) de todos los coeficientes señala un valor de 0,000. La significatividad del estadístico de Hosmer y Lemeshow es 0,117. Y el R² de Nagelkerke alcanza un valor de 0,496. Los resultados denotan, pues, un ajuste aceptable.

Igualmente, la matriz de clasificación indica que se han clasificado correctamente el 77,0% de las empresas, obteniéndose un mayor nivel de predicción en el caso de las empresas insolventes, con un 78,7% de acierto, y clasificando correctamente el 75,4% de las empresas solventes. Por su parte, el nivel de clasificación en la muestra externa, utilizada para comprobar la capacidad de generalización del modelo, ha resultado ser del 74,5%.

4.3.5 RESULTADOS MODELO COMERCIO Y SERVICIOS

4.3.5.1 Análisis Exploratorio

En este apartado se analizan los principales parámetros estadísticos de las variables correspondientes a la muestra del sector comercio y servicios.

4.3.5.1.1 Análisis Descriptivo

Observando las medias de las variables se comprueba que, en la muestra M.II, sólo VMG1 y VMSCS3 presentan diferente signo según se trate de empresas solventes e insolventes. Además, las variables VMG3, VMG7, VMG9, y VMSCS4 presentan medias inferiores en el caso de las empresas solventes. Y respecto a la desviación típica, encontramos valores reducidos, salvo para VMSCS2 y VMSCS4 (tabla 4.57).

**TABLA 4.57 ESTADÍSTICOS DESCRIPTIVOS
(MODELO COMERCIO Y SERVICIOS, M.II)**

		MEDIA		MEDIANA		DESV. TÍPICA		MINIMO		MAXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,01	-0,03	0,01	0,00	0,04	0,09	-0,22	-0,76	0,19	0,19
	VMG2	1,52	1,18	1,16	1,10	1,27	0,59	0,19	0,01	10,19	4,26
	VMG3	0,30	0,40	0,30	0,40	0,28	0,28	-0,39	-0,29	0,92	0,97
	VMG4	0,04	0,00	0,03	0,01	0,06	0,08	-0,28	-0,38	0,24	0,18
	VMG5	1,70	1,46	1,35	1,15	1,30	1,24	0,10	0,01	8,09	7,92
	VMG6	0,88	0,70	0,65	0,66	0,94	0,48	0,00	0,01	7,07	2,61
	VMG7	0,21	0,32	0,16	0,30	0,21	0,25	0,00	0,00	0,81	1,32
	VMG8	0,69	0,66	0,77	0,73	0,24	0,26	0,03	0,01	1,00	1,00
	VMG9	0,08	0,28	0,04	0,02	0,32	1,39	-1,31	-4,44	2,28	8,20
VARIABLES MODELO SECTORIAL (COMERCIO Y SERVICIOS)	VMSCS1	0,06	0,01	0,05	0,03	0,06	0,09	-0,22	-0,35	0,26	0,40
	VMSCS2	4,78	0,91	2,00	0,93	30,88	20,60	-227,6	-40,61	174,93	229,93
	VMSCS3	0,09	-0,01	0,07	0,02	0,16	0,21	-0,41	-1,07	1,22	1,08
	VMSCS4	13,03	28,61	6,15	6,04	19,48	93,34	0,20	0,08	116,57	823,15
	VMSCS5	1,70	1,46	1,35	1,15	1,30	1,24	0,10	0,01	8,09	7,92

4.3.5.1.2 Análisis de Normalidad

La tabla 4.58 muestra los resultados del test de Kolmogorov-Smirnov para las empresas solventes e insolventes pertenecientes al sector comercio y servicios, a un nivel de significación del 5%. Con la excepción de VMG3, y para empresas solventes e insolventes, todas las variables en M.II, tienen un nivel de significación inferior al 5%. Esto implica rechazar la hipótesis nula de normalidad.

**TABLA 4.58 ANÁLISIS DE NORMALIDAD
(MODELO COMERCIO Y SERVICIOS, M.II)**

		M.II (2 AÑOS ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	0,230	0,000	0,284	0,000
	VMG2	0,287	0,000	0,208	0,000
	VMG3	0,056	0,200	0,045	0,200
	VMG4	0,194	0,000	0,271	0,000
	VMG5	0,214	0,000	0,181	0,000
	VMG6	0,299	0,000	0,206	0,000
	VMG7	0,213	0,000	0,097	0,000
	VMG8	0,124	0,000	0,098	0,000
	VMG9	0,420	0,000	0,305	0,000
VARIABLES MODELO SECTORIAL (COMERCIO Y SERVICIOS)	VMSCS1	0,150	0,000	0,241	0,000
	VMSCS2	0,317	0,000	0,367	0,000
	VMSCS3	0,269	0,000	0,288	0,000
	VMSCS4	0,365	0,000	0,451	0,000
	VMSCS5	0,214	0,000	0,181	0,000

4.3.5.1.3 Tests no paramétricos

Tan solo VMG9 y VMSCS4 registran un test de Mann-Whitney con un nivel de significación superior al 5% ($p > 0,05$). Asimismo, VMG5, VMSCS4 y VMSCS5 presentan un test de Kolmogorov-Smirnov con un p-valor superior a 0,05. Con las excepciones apuntadas, se rechazaría la hipótesis nula de igualdad en las distribuciones (tabla 4.59).

**TABLA 4.59 TESTS NO PARAMÉTRICOS
(MODELO COMERCIO Y SERVICIOS, M.II)**

		M.II (2 AÑOS ANTES DE LA QUIEBRA)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	-8,060	0,000	3,606	0,000
	VMG2	-3,385	0,001	1,587	0,013
	VMG3	3,616	0,000	2,196	0,000
	VMG4	-5,718	0,000	2,685	0,000
	VMG5	-2,285	0,022	1,288	0,072
	VMG6	-2,017	0,044	1,609	0,011
	VMG7	6,721	0,000	3,416	0,000
	VMG8	-2,737	0,006	1,544	0,017
	VMG9	-1,782	0,075	1,616	0,011
VARIABLES MODELO SECTORIAL (COMERCIO Y SERVICIOS)	VMSCS1	-5,663	0,000	2,370	0,000
	VMSCS2	-8,468	0,000	3,940	0,000
	VMSCS3	-6,256	0,000	2,753	0,000
	VMSCS4	-0,333	0,739	0,793	0,555
	VMSCS5	-2,285	0,022	1,288	0,072

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

4.3.5.1.4 Análisis de Correlación

Para concluir el análisis exploratorio del sector comercio y servicios en M.II se procederá al cálculo de las correlaciones que se producen entre las variables. Para ello, en la tabla 4.60 se ofrecen los coeficientes de correlación de Pearson. Se observa una elevada correlación entre los pares de variables VMG1 y VMG4 (0,976), VMG2 y VMG6 (0,952), VMG4 y VMSCS1 (0,987), VMG5 y VMSCS5 (1,000) y VMSCS1 y VMG1 (0,963).

**TABLA 4.60 ANÁLISIS DE CORRELACIÓN BILATERAL
(MODELO COMERCIO Y SERVICIOS, M.II)**

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	VMSCS1	VMSCS2	VMSCS3	VMSCS4	VMSCS5	INSOLV
VMG1	1	0,069	0,049	0,976	-0,174	0,088	-0,141	0,112	0,091	0,963	0,518	0,608	-0,028	-0,174	-0,202
VMG2	0,069	1	0,030	0,047	-0,104	0,952	-0,154	0,129	-0,019	0,029	0,080	0,241	-0,047	-0,104	-0,148
VMG3	0,049	0,030	1	0,057	-0,103	-0,029	0,022	0,467	-0,039	0,007	0,032	-0,045	-0,034	-0,103	0,159
VMG4	0,976	0,047	0,057	1	-0,141	0,069	-0,084	0,080	0,107	0,987	0,499	0,629	-0,010	-0,141	-0,161
VMG5	-0,174	-0,104	-0,103	-0,141	1	-0,055	-0,187	0,080	0,015	-0,107	0,041	-0,003	0,065	1,000	-0,114
VMG6	0,088	0,952	-0,029	0,069	-0,055	1	-0,146	0,146	-0,005	0,055	0,122	0,312	-0,012	-0,055	-0,125
VMG7	-0,141	-0,154	0,022	-0,084	-0,187	-0,146	1	-0,329	-0,068	-0,086	-0,154	-0,113	0,009	-0,187	0,311
VMG8	0,112	0,129	0,467	0,080	0,080	0,146	-0,329	1	0,049	0,028	0,031	-0,052	-0,077	0,080	-0,121
VMG9	0,091	-0,019	-0,039	0,107	0,015	-0,005	-0,068	0,049	1	0,107	0,157	0,072	0,000	0,015	0,072
VMSCS1	0,963	0,029	0,007	0,987	-0,107	0,055	-0,086	0,028	0,107	1	0,486	0,630	0,025	-0,107	-0,159
VMSCS2	0,518	0,080	0,032	0,499	0,041	0,12	-0,154	0,031	0,157	0,486	1	0,449	0,000	0,041	-0,142
VMSCS3	0,608	0,241	-0,045	0,629	-0,003	0,312	-0,113	-0,052	0,072	0,630	0,449	1	0,001	-0,003	-0,181
VMSCS4	-0,028	-0,047	-0,034	-0,010	0,065	-0,012	0,009	-0,077	0,000	0,025	0,000	0,001	1	0,065	0,063
VMSCS5	-0,174	-0,104	-0,103	-0,141	1,000	-0,055	-0,187	0,080	0,015	-0,107	0,041	-0,003	0,065	1	-0,114
INSOLV	-0,202	-0,148	0,159	-0,161	-0,114	-0,125	0,311	-0,121	0,072	-0,159	-0,142	-0,181	0,063	-0,114	1

4.3.5.2 Análisis Confirmatorio

Con objeto de confirmar los resultados del análisis exploratorio precedente, se ha estimado el modelo Logit que aparece a continuación.

4.3.5.2.1 Modelo Logit Sector Comercio y Servicios

Los resultados del Logit estimado para la muestra de empresas del sector comercio y servicios (en M.II) aparecen en la tabla 4.61. Dicho modelo está integrado por las variables VMG1, VMG2, VMG3 y VMG7.

Con respecto al signo de los coeficientes obtenidos para cada una de estas variables se observa que las variables VMG1 y VMG2 presentan signo negativo, lo cual denota una relación inversa con la insolvencia. Por su parte, las variables VMG3 y VMG7 presentan signos positivos.

Destaca la variable VMG7 con un Odds Ratio de 8,314, lo que significa que cuando el ratio Total Deuda/Total Activo aumenta en un punto, la probabilidad de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 8,314 veces.

Con respecto a la robustez y precisión del modelo en su conjunto, la prueba de Razón de Verosimilitud (RV) de todos los coeficientes arroja un valor su significatividad de 0,000. La precisión de clasificación del modelo es del

79,5%. Y el test de Hosmer y Lemeshow muestra una significatividad del 0,374.

TABLA 4.61 MODELO LOGIT
SECTOR COMERCIO Y SERVICIOS. M.II

Ajuste del modelo		Valor	
Significatividad prueba Ómnibus		0,000	
Significatividad Test de Hosmer y Lemeshow		0,374	
-2 log verosimilitud (-2LL)		364,380	
R ² Cox & Snell		0,208	
R ² Nagelkerke		0,277	
Matriz de clasificación		%	
Dentro de la muestra		70,9	
	Insolventes	70,9	
	Solventes	70,9	
Muestra de testeo		79,5	
	Insolventes	79,4	
	Solventes	79,7	
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-9,135	0,000	0,000
VMG2	-0,307	0,735	0,027
VMG3	1,637	5,138	0,000
VMG7	2,118	8,314	0,000
Constante	-0,788	0,455	0,012

Notas: VMG1 = Beneficio después de Impuestos/Total Activo; VMG2 = Activo corriente/Pasivo corriente; VMG3 = Fondo de Maniobra/Total Activo; VMG7 = Total Deuda/Total Activo.

4.3.6 RESULTADOS SECTOR HOSTELERÍA

4.3.6.1 Análisis Exploratorio

En este epígrafe se aborda el análisis exploratorio de las variables correspondientes al sector hostelería en M.II.

4.3.6.1.1 Análisis Descriptivo

Conforme al análisis descriptivo realizado, se comprueba que VMG1, VMG4, y VMSH2 presentan valores medios con diferente signo según se trate de empresas solventes e insolventes (tabla 4.62).

Se aprecia, además, cómo el 80% de los valores medios en las empresas solventes supera al de las empresas insolventes. Para el caso de la mediana, este porcentaje es del 90%.

En particular, sólo la variable VMG7 presenta una media inferior en el caso de las empresas solventes frente a las insolventes.

La desviación típica de las variables también presenta valores moderados, salvo para VMSH5.

**TABLA 4.62 ESTADÍSTICOS DESCRIPTIVOS
(MODELO HOSTELERÍA, M.II)**

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
VARIABLES MODELO GLOBAL	VMG1	0,03	-0,16	0,02	-0,08	0,06	0,22	-0,44	-0,92	0,26	0,08
	VMG2	1,61	0,68	1,07	0,44	1,73	0,73	0,00	0,02	8,36	3,68
	VMG3	0,13	0,06	0,07	0,04	0,22	0,20	-0,38	-0,38	0,79	0,74
	VMG4	0,06	-0,15	0,04	-0,07	0,07	0,22	-0,16	-0,87	0,35	-0,12
	VMG5	1,95	1,25	1,59	0,89	1,52	1,21	0,00	0,00	7,39	5,84
	VMG6	1,14	0,54	0,64	0,28	1,49	0,66	0,00	0,00	7,71	3,45
	VMG7	0,42	0,56	0,36	0,53	0,35	0,35	0,00	0,00	2,33	1,67
	VMG8	0,40	0,32	0,36	0,26	0,27	0,26	0,00	0,00	0,99	1,00
	VMG9	0,21	-0,00	0,08	0,08	0,50	1,54	-0,67	-9,91	3,14	4,23
VARIABLES MODELO SECTORIAL (HOSTELERÍA)	VMSH1	0,49	-0,13	0,37	-0,09	0,51	0,43	-0,48	-2,13	3,08	1,03
	VMSH2	0,11	-0,10	0,09	-0,03	0,08	0,21	-0,16	-0,85	0,40	0,23
	VMSH3	6,13	-5,37	2,99	-1,26	9,90	23,29	-25,23	-105,66	51,05	61,64
	VMSH4	3,71	1,45	0,84	0,08	8,59	9,80	-8,35	-47,34	44,74	32,33
	VMSH5	250,16	152,45	38,95	12,12	806,83	608,31	0,12	0,01	6435,7	5357,0
	VMSH6	1,16	-0,12	0,21	-0,08	3,50	2,32	-0,18	-12,61	22,19	19,53

4.3.6.1.2 Análisis de Normalidad

Todas las variables presentan un nivel de significación inferior al 5%, lo que implica rechazar la hipótesis nula de normalidad (tabla 4.63).

**TABLA 4.63 ANÁLISIS DE NORMALIDAD
(MODELO HOSTELERÍA, M.II)**

		M.II (2 AÑOS ANTES DE LA QUIEBRA)			
		SOLVENTES		INSOLVENTES	
		K-S	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	0,369	0,000	0,328	0,000
	VMG2	0,409	0,000	0,297	0,000
	VMG3	0,143	0,000	0,157	0,000
	VMG4	0,335	0,000	0,296	0,000
	VMG5	0,427	0,000	0,325	0,000
	VMG6	0,418	0,000	0,247	0,000
	VMG7	0,220	0,000	0,167	0,000
	VMG8	0,090	0,000	0,120	0,000
	VMG9	0,345	0,000	0,331	0,000
VARIABLES MODELO SECTORIAL (HOSTELERÍA)	VMSH1	0,383	0,000	0,154	0,000
	VMSH2	0,309	0,000	0,296	0,000
	VMSH3	0,344	0,000	0,230	0,000
	VMSH4	0,398	0,000	0,317	0,000
	VMSH5	0,512	0,000	0,505	0,000
	VMSH6	0,367	0,000	0,351	0,000

NOTA: K-S: test de Kolmogorov-Smirnov

4.3.6.1.3 Tests no paramétricos

Conforme a los resultados que aparecen en la tabla 4.64, tan solo VMG3 presenta un test de Mann-Whitney y de Kolmogorov-Smirnov con un nivel de significación superior al 5% ($p > 0,05$). Por tanto, y con la excepción apuntada, se rechaza la hipótesis nula de igualdad en las distribuciones en todas las variables consideradas en el sector hostelería.

**TABLA 4.64 TESTS NO PARAMÉTRICOS
(MODELO HOSTELERÍA, M.II)**

		M.II (2 AÑOS ANTES)			
		Test de Mann-Whitney		Test de Kolmogorov-Smirnov	
		M-W	Sig.	K-S	Sig.
VARIABLES MODELO GLOBAL	VMG1	-11,758	0,000	6,251	0,000
	VMG2	-5,657	0,000	2,716	0,000
	VMG3	-1,155	0,248	1,000	0,270
	VMG4	-10,367	0,000	5,480	0,000
	VMG5	-4,551	0,000	2,204	0,000
	VMG6	-4,165	0,000	2,054	0,000
	VMG7	3,810	0,000	1,886	0,002
	VMG8	-4,363	0,000	2,181	0,000
	VMG9	-1,868	0,062	2,724	0,000
VARIABLES MODELO SECTORIAL (HOSTELERÍA)	VMSH1	-9,489	0,000	4,419	0,000
	VMSH2	-9,333	0,000	4,466	0,000
	VMSH3	-5,129	0,000	3,453	0,000
	VMSH4	-1,659	0,097	1,854	0,002
	VMSH5	-6,667	0,000	3,437	0,000
	VMSH6	-9,249	0,000	5,001	0,000

NOTA: K-S: test de Kolmogorov-Smirnov; M-W: test de Mann-Whitney

4.3.6.1.4 Análisis de Correlación

Se ha comprobado que se registra una elevada correlación entre los siguientes pares de variables: VMG1 y VMG4 (0,965), VMG2 y VMG6 (0,988), VMG1 y VMSH2 (0,909) y VMG4 y VMSH2 (0,980) (tabla 4.65).

**TABLA 4.65 ANÁLISIS DE CORRELACIÓN BILATERAL
(MODELO HOSTELERÍA, M.II)**

	VMG1	VMG2	VMG3	VMG4	VMG5	VMG6	VMG7	VMG8	VMG9	VMSH1	VMSH2	VMSH3	VMSH4	VMSH5	VMSH6	INSOLV
VMG1	1	0,045	0,218	0,965	0,210	0,044	-0,009	0,016	0,011	0,078	0,909	0,039	0,013	0,003	0,626	-0,202
VMG2	0,045	1	0,065	0,029	-0,016	0,988	-0,057	0,134	0,017	0,455	0,023	-0,017	-0,005	-0,011	0,100	-0,123
VMG3	0,218	0,065	1	0,166	-0,069	0,034	-0,063	0,368	-0,075	0,003	0,120	-0,183	-0,026	-0,043	0,066	-0,064
VMG4	0,965	0,029	0,166	1	0,276	0,027	0,007	0,071	0,048	0,123	0,980	0,039	0,013	0,000	0,607	-0,231
VMG5	0,210	-0,016	-0,069	0,276	1	-0,015	-0,065	0,144	-0,101	0,033	0,331	-0,037	-0,020	0,183	-0,473	-0,057
VMG6	0,044	0,988	0,034	0,027	-0,015	1	-0,087	0,121	0,017	0,432	0,021	-0,023	-0,018	-0,011	0,188	-0,117
VMG7	-0,009	-0,057	-0,063	0,007	-0,065	-0,087	1	-0,271	0,093	-0,039	-0,008	0,209	0,131	-0,063	-0,169	0,161
VMG8	0,016	0,134	0,368	0,071	0,144	0,121	-0,271	1	-0,087	0,102	0,076	-0,177	-0,105	-0,020	0,137	-0,211
VMG9	0,011	0,017	-0,075	0,048	-0,101	0,017	0,093	-0,087	1	0,041	0,048	0,055	0,259	-0,002	0,018	-0,102
VMSH1	0,078	0,455	0,003	0,123	0,033	0,432	-0,039	0,102	0,041	1	0,135	0,007	0,008	-0,006	0,235	-0,138
VMSH2	0,909	0,023	0,120	0,980	0,331	0,021	-0,008	0,076	0,048	0,135	1	0,026	0,008	0,021	0,603	-0,237
VMSH3	0,039	-0,017	-0,183	0,039	-0,037	-0,023	0,209	-0,177	0,055	0,007	0,026	1	0,208	-0,019	-0,031	-0,086
VMSH4	0,013	-0,005	-0,026	0,013	-0,020	-0,018	0,131	-0,105	0,259	0,008	0,008	0,208	1	-0,009	-0,020	-0,021
VMSH5	0,003	-0,011	-0,043	0,000	0,183	-0,011	-0,063	-0,020	-0,002	-0,006	0,021	-0,019	-0,009	1	0,063	-0,030
VMSH6	0,626	0,100	0,066	0,607	-0,473	0,188	-0,169	0,137	0,018	0,235	0,603	-0,031	-0,020	0,063	1	-0,174
INSOLV	-0,202	-0,123	-0,064	-0,231	-0,057	-0,117	0,161	-0,211	-0,102	-0,138	-0,237	-0,086	-0,021	-0,030	-0,174	1

4.3.6.2 Análisis Confirmatorio

A continuación aparecen los resultados del modelo logit estimado para el sector hostelería según M.II.

4.3.6.2.1 Modelo Logit Sector Hostelería

Como en el resto de los modelos estimados, se aplicará el método de máximo verosimilitud, de forma que se irán realizando sucesivas iteraciones hasta obtener la estimación más probable. En la tabla 4.66 se muestran los resultados obtenidos.

Las variables que han sido seleccionadas por el modelo han sido VMG2, VMG4 y VMG6. Además, se puede apreciar que VMG2 y VMG4 presentan signo negativo, lo cual denota una relación inversa con la insolvencia. Por el contrario, VMG6 presenta un signo positivo, evidenciando que cuanto más alto es este ratio mayor es la probabilidad de que la empresa entre en quiebra.

En el caso de VMG6, al tener un Odds Ratio de 2,080 indica que cuando el ratio Quick aumenta en un punto, la probabilidad de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 2,080 veces.

Finamente, también es destacable que tanto la prueba Ómnibus como la matriz de clasificación (se han clasificado correctamente el 81,0% de las empresas) y el test de Hosmer y Lemeshow apuntan a un ajuste adecuado del modelo.

**TABLA 4.66 MODELO LOGIT
SECTOR HOSTELERÍA. M.II**

Ajuste del modelo		Valor	
Significatividad prueba Ómnibus		0,000	
Significatividad Test de Hosmer y Lemeshow		0,082	
-2 log verosimilitud (-2LL)		234,212	
R ² Cox & Snell		0,284	
R ² Nagelkerke		0,379	
Matriz de clasificación		%	
Dentro de la muestra		72,6	
	Insolventes	69,8	
	Solventes	75,7	
Muestra de testeo		81,0	
	Insolventes	79,1	
	Solventes	82,9	
Variables	Coefficiente	Odds ratio	Significatividad
VMG2	-0,476	0,621	0,058
VMG4	-14,835	0,000	0,000
VMG6	0,733	2,080	0,010
Constante	0,231	1,259	0,317

Notas: VMG2 = Activo Corriente/Pasivo Corriente; VMG4 = EBIT/Total Activo;
VMG6 = Quick Ratio.

4.3.7 COMPARACIÓN DEL MODELO GLOBAL CON LOS MODELOS SECTORIALES EN M.II.

Una vez obtenidos los modelos de regresión para cada uno de los sectores objeto de este estudio y para el modelo global (obtenido con una muestra compuesta por empresas de todos los sectores), procede a dar respuesta a la hipótesis H_1 formulada en el capítulo primero de esta tesis, en referencia a M.II. Esta hipótesis plantea si un

modelo global era igual o distinto a los modelos sectoriales estimados a partir de muestras específicas de cada sector.

A tal efecto, y conforme a los resultados que aparecen en la tabla 4.67, es posible concluir que los modelos estimados son diferentes, pues presentan regresores distintos. Por consiguiente, se acepta la hipótesis H_1 planteada.

TABLA 4.67 MODELOS GLOBAL Y SECTORIALES OBTENIDOS PARA M.II

Ajuste del Modelo	Global	Agricultura	Industria	Construcción	Comercio y Servicios	Hostelería
Prueba Ómnibus	0,000	0,000	0,000	0,000	0,000	0,000
Test Hosmer-Lemeshow	0,100	0,210	0,189	0,117	0,374	0,082
R ² Nagelkerke	0,365	0,327	0,832	0,496	0,277	0,379
Matriz de Clasificación (%)						
Fuera de la Muestra						
Solventes	74,3	45,0	83,3	76,6	79,7	82,9
Insolventes	71,1	88,5	72,5	72,1	79,4	79,1
Total	72,7	69,6	78,1	74,5	79,5	81,0
Regresores	VMG1 VMG2 VMG3 VMG9	VMG1 VMSA1	VMG1	VMG1 VMG3 VMG6 VMG7 VMG8 VMSI1	VMG1 VMG2 VMG3 VMG7	VMG2 VMG4 VMG6
			VMG2			
			VMG3			
			VMG4			
			VMG6			
			VMG7			
			VMG8			
			VMG9			
			VMSI4			

4.3.8 MODELO LOGIT GLOBAL CON VARIABLES CUALITATIVAS SECTORIALES

En este apartado se aborda el contraste de la hipótesis H_2 sobre si la inclusión de variables cualitativas (representativas de los sectores) en el modelo global, mejora el ajuste y la precisión del mismo.

Para dar respuesta a esta cuestión, se abordará, en primer lugar, la estimación de un modelo global elaborado con una muestra de todos los sectores económicos para M.II introduciendo variables dummies representativas de dichos sectores. En segundo lugar, se compararan los resultados así obtenidos con los correspondientes al modelo global estimado sin la participación de esas variables cualitativas.

Los resultados del modelo Logit global con dummies estimado para M.II aparecen en la tabla 4.68. Las variables seleccionadas han sido VMG1, VMG2, VMG3, VMG9, y la dummy del sector industria.

Al analizar los signos de los coeficientes obtenidos para cada una de estas variables se puede apreciar que todas las variables excepto VMG3 presentan signo negativo, además del coeficiente de la dummy del sector industria. Esto denota una relación inversa con la insolvencia. Por lo tanto, cuanto más alto son estos ratios menor es la probabilidad de que la empresa entre en quiebra. Por su parte, la variable VMG3 presenta signo positivo en su coeficiente. En este caso, la relación con la

insolvencia es directa, es decir, cuanto más alto es el ratio, mayor es la probabilidad de insolvencia.

Si se analizan los Odds ratio de las variables, como era de suponer, tan solo son mayores que 1 el correspondiente a VMG3 en tanto que es la única variable con signo positivo, siendo valores menores a 1 en el resto de los casos.

En el caso de VMG3 al tener signo positivo y un Odds Ratio de 4,470 indica que cuando el ratio Fondo de Maniobra/Total Activo aumenta en un punto, la probabilidad de que la empresa sea insolvente, respecto a la probabilidad de que no lo sea, aumenta 4,470 veces.

Por lo que se refiere a la significatividad de las variables incluidas en el modelo, todas las variables presentan niveles de significatividad por debajo del 5%, salvo la dummy del sector industria, que es significativa al 10%.

Una vez valorada la significación de cada variable de forma individual, procede analizar distintos tests para constatar la robustez y precisión del modelo en su conjunto.

Por lo que se refiere a la prueba de Razón de Verosimilitud (RV) de todos los coeficientes, el valor obtenido es de 0,000, concluyéndose que el ajuste es

adecuado, dado que al menos uno de los predictores está significativamente relacionado con la variable dependiente.

TABLA 4.68 MODELO LOGIT GLOBAL CON VARIABLES CUALITATIVAS. M.II

Ajuste del modelo		Valor	
Significatividad prueba Ómnibus		0,000	
Significatividad Test de Hosmer y Lemeshow		0,100	
-2 log verosimilitud (-2LL)		450,885	
R ² Cox & Snell		0,274	
R ² Nagelkerke		0,365	
Matriz de clasificación		%	
Dentro de la muestra		73,3	
	Insolventes	70,8	
	Solventes	75,8	
Muestra de testeo		72,7	
	Insolventes	71,1	
	Solventes	74,3	
Variables	Coeficiente	Odds ratio	Significatividad
VMG1	-14,530	0,000	0,000
VMG2	-0,086	0,918	0,024
VMG3	1,497	4,470	0,000
VMG9	-0,425	0,653	0,009
Dummy (Industria)	-0,470	0,625	0,083
Constante	-0,048	0,953	0,800

Notas: VMG1 = Beneficio después de Impuestos/Total Activo; VMG2 = Activo corriente/Pasivo corriente; VMG3 = Fondo de Maniobra/Total Activo; VMG9 = Beneficio después de Impuestos/Patrimonio Neto.

En segundo lugar, el análisis de la capacidad predictiva del modelo mediante la matriz de clasificación nos indica que se han clasificado correctamente el 73,3% de las empresas, obteniéndose un mayor nivel de

clasificación en el caso de las empresas solventes, con un 75,8% de acierto. Con respecto al nivel de clasificación en la muestra externa, este porcentaje ha resultado ser del 72,7%.

La evaluación de la bondad del ajuste del modelo a partir del test de Hosmer y Lemeshow indica que no existen diferencias significativas entre las clasificaciones observadas y predichas.

Finalmente, el coeficiente de Cox y Snell asciende a 0,274 y el R^2 de Nagelkerke alcanza un valor de 0,365.

Una vez estimado y analizado el modelo global con variables cualitativas, procede comparar los resultados obtenidos con el modelo global sin la inclusión de dichas variables de tipo cualitativo.

La comparación va a permitir dar respuesta a la segunda hipótesis, sobre la superioridad de un modelo global sin dummies sectoriales, frente a un modelo que incluyera este tipo de variables representativa de los distintos sectores.

Para llevar a cabo esta comparación se han utilizado tres criterios de selección de modelos, ya tratados en el capítulo 2. Estos criterios son los de Akaike, el bayesiano de Schwartz y el criterio de Hannann-Quinn. La regla de decisión para los tres criterios es la misma: el mejor modelo es el que ofrece el menor valor en la comparación.

De acuerdo con esta regla, y como puede comprobarse a partir de los valores que aparecen en la tabla 4.69, el modelo global con dummies es superior al modelo global estimado sin la participación de variables cualitativas sectoriales, en M.II. De esta forma, los resultados obtenidos permiten aceptar la hipótesis H_2 implicando que la inclusión de variables cualitativas representativas de sectores económicos enriquece y aumenta la capacidad de explicación del modelo global.

**TABLA 4.69 COMPARACIÓN DE MODELOS GLOBALES.
M.II.**

Ajuste del Modelo	Global SIN DUMMYS	Global CON DUMMYS
Prueba Ómnibus	0,000	0,000
Test Hosmer-Lemeshow	0,139	0,100
R ² Nagelkerke	0,359	0,365
Matriz de Clasificación (%)		
Fuera de la Muestra		
Solventes	72,3	74,3
Insolventes	69,9	71,1
Total	71,1	72,7
Criterio de Akaike	461,92	458,89
Criterio de Schwartz	468,52	466,19
Criterio de Hannan-Quinn	458,42	455,55

4.3.9 ANÁLISIS DE ROBUSTEZ

En este apartado se llevará a cabo un análisis de la robustez de los resultados obtenidos en los diferentes modelos, tal y como se efectuó también para la muestra M.II. Este análisis constará de dos partes: por un lado, ofrecerá los resultados de una validación cruzada de los modelos, comprobando la capacidad del modelo global con variables cualitativas para producir buenos resultados utilizando las muestras de cada uno de los sectores. Por otro lado, se aplicarán los criterios de Akaike, de Schwartz y de Hanann-Quinn, todos ellos criterios informativos de selección de modelos.

Este análisis de robustez permitirá conocer si se acepta o se rechaza la hipótesis H_3 planteada, y que hace referencia a la capacidad del modelo global con dummies para predecir correctamente con las muestras de los diferentes sectores.

De los resultados que aparecen en las tablas 4.70 y 4.71, donde se lleva a cabo la validación cruzada del modelo global con dummies para M.II en las distintas muestras sectoriales, se deduce que dicho modelo global es capaz de predecir con éxito sobre las distintas muestras sectoriales. Estos resultados permiten aceptar la hipótesis H_3 .

TABLA 4.70 RESULTADOS DE LOS MODELOS DE REGRESIÓN LOGÍSTICA PARA M.II.

Ajuste del Modelo	Global	Agricultura	Industria	Construcción	Comercio y Servicios	Hostelería
Prueba Ómnibus	0,000	0,000	0,000	0,000	0,000	0,000
Test Hosmer-Lemeshow	0,100	0,210	0,189	0,117	0,374	0,082
R ² Nagelkerke	0,365	0,327	0,832	0,496	0,277	0,379
Matriz de Clasificación (%)						
Fuera de la Muestra						
Solventes	74,3	45,0	83,3	76,6	79,7	82,9
Insolventes	71,1	88,5	72,5	72,1	79,4	79,1
Total	72,7	69,6	78,1	74,5	79,5	81,0
Criterio de Akaike	729,50	102,33	84,26	373,28	373,38	240,21
Criterio de Schwartz	736,19	104,75	101,17	383,70	379,58	245,68
Criterio de Hannann-Quinn	726,16	100,36	76,86	368,15	369,02	237,71

TABLA 4.71 ROBUSTEZ REGRESIÓN LOGÍSTICA DEL MODELO GLOBAL CON VARIABLES CUALITATIVAS. M.II.

Ajuste del Modelo	Agricultura	Industria	Construcción	Comercio y Servicios	Hostelería
Prueba Ómnibus	0,004	0,000	0,000	0,000	0,000
Test Hosmer-Lemeshow	0,578	0,402	0,654	0,600	0,100
R ² Nagelkerke	0,369	0,415	0,374	0,461	0,605
Matriz de Clasificación (%)					
Fuera de la Muestra					
Solventes	69,2	79,3	72,0	75,9	76,6
Insolventes	63,6	72,2	69,4	74,7	84,0
Total	66,7	75,9	70,8	75,3	80,4
Criterio de Akaike	58,72	175,93	170,28	167,73	131,49
Criterio de Schwartz	63,55	183,45	177,23	174,93	138,77
Criterio de Hannann-Quinn	54,77	172,64	166,86	164,37	128,14

4.4 ANÁLISIS COMPARATIVO DE RESULTADOS

En la tabla 4.72 aparece los resultados de la comparación de regresores entre los modelos globales y los modelos sectoriales correspondientes a M.I y M.II. Conforme a dichos resultados se comprueba que los modelos globales y sectoriales tanto en M.I como en M.II son distintos. Es decir, aunque pueden contener regresores comunes en algunos casos, no son modelos iguales. Por tanto, queda aceptada definitivamente la Hipótesis H_1 sobre las diferencias en ambos tipos de modelos.

En relación con la hipótesis H_2 , sobre la introducción de variables cualitativas sectoriales en los modelos globales, la tabla 4.73 ofrece una comparación de los resultados obtenidos tanto para M.I como para M.II. Como se puede comprobar, la inclusión de variables dummies mejora los modelos globales estimados a tenor de la comparación efectuada mediante los criterios de Akaike, Schwartz y Hannann-Quinn.

TABLA 4.72 COMPARACIÓN ENTRE LOS MODELOS GLOBALES Y SECTORIALES EN M.I Y M.II

Ajuste del Modelo	Global	Agricultura	Industria	Construcción	Comercio y Servicios	Hostelería
Regresores de los modelos en M.I.			VMG1			
			VMG2		VMG1	VMG1
	VMG1	VMG1	VMG3	VMG1	VMG2	VMG2
	VMG3	VMG6	VMG4	VMG6	VMG3	VMG5
	VMG4	VMSA2	VMG5	VMG7	VMG7	
	VMG5		VMG6	VMG8	VMG8	VMSH1
			VMG8	VMSC1	VMSCS4	VMSH3
			VMSI7			
Regresores de los modelos en M.II.			VMG1			
			VMG2	VMG1		
	VMG1		VMG3	VMG3	VMG1	VMG2
	VMG2	VMG1	VMG4	VMG6	VMG2	VMG4
	VMG3	VMSA1	VMG6	VMG7	VMG3	VMG6
	VMG9		VMG7	VMG8	VMG7	
			VMG8	VMSC1		
			VMG9			
			VMSI4			

TABLA 4.73 COMPARACIÓN ENTRE LOS MODELOS GLOBALES EN M.I Y M.II.

Criterios de Selección de Modelos	M.I		M.II	
	Global	Global	Global	Global
	SIN DUMMYS	CON DUMMYS	SIN DUMMYS	CON DUMMYS
Criterio de Akaike (AIC)	398,74	330,40	461,92	458,89
Criterio de Schwartz (BIC)	405,34	341,35	468,52	466,19
Criterio de Hannan-Quinn (HQC)	395,25	325,39	458,42	455,55

Finalmente, y en relación a la hipótesis H_3 , relativa a si el modelo global era capaz de predecir correctamente con las muestras de los diferentes sectores, los resultados que aparecen en la tabla 4.74 indican que esta hipótesis es aceptada solo parcialmente. Se ha podido comprobar que el modelo del sector industria en M.II es superior al modelo global estimado.

TABLA 4.74 RESULTADOS DE LA COMPARACIÓN ENTRE MODELOS GLOBALES Y SECTORIALES

	1 AÑO ANTES QUIEBRA (M.I)			2 AÑOS ANTES QUIEBRA (M.II)		
	<i>AIC</i>	<i>BIC</i>	<i>HQC</i>	<i>AIC</i>	<i>BIC</i>	<i>HQC</i>
AGRICULTURA						
<i>Modelo</i>	84,31	87,96	81,36	102,33	104,75	100,36
<i>Validación Cruzada</i>	56,34	63,64	50,43	58,72	63,55	54,77
INDUSTRIA						
<i>Modelo</i>	175,94	190,97	169,36	84,26	101,17	76,86
<i>Validación Cruzada</i>	112,69	123,96	107,75	175,93	183,45	172,64
CONSTRUCCIÓN						
<i>Modelo</i>	270,01	278,69	265,73	373,28	383,70	368,15
<i>Validación Cruzada</i>	128,25	138,68	123,12	170,28	177,23	166,86
COMERCIO Y SERVICIOS						
<i>Modelo</i>	229,61	240,41	224,56	373,38	379,58	369,02
<i>Validación Cruzada</i>	117,10	127,90	112,05	167,73	174,93	164,37
HOSTELERÍA						
<i>Modelo</i>	259,87	268,98	255,70	240,21	245,68	237,71
<i>Validación Cruzada</i>	98,44	109,38	93,43	131,49	138,77	128,14

BIBLIOGRAFÍA CAPÍTULO 4

Altman, D.G. (1991). *Practical Statistics for Medical Research*. Londres. Chapman.

Fernández, M.A.; Gil, A.M. y P. López (2009). *Estructura y Equilibrio financiero de las PYMES Españolas ante las Nuevas Exigencias de Información Financiera*. Fundación de las Cajas de Ahorros. Madrid.

Hair, J. (2008). *Análisis Multivariante*. Editorial Prentice Hall. Madrid.

Hair, J.F., Anderson R.E., Tatham, R.L. y W.C. Black (1999). *Análisis multivariante*, 5ª edición. Editorial Prentice Hall. Madrid.

Kolmogorov, A.N. (1957). On the representations of continuous functions of many variables by superpositions of continuous functions of one variable and addition. *Doklady Akademii Nauk USSR*, 114:5, 953-956.

Levy, J.P. y J. Varela (2003). *Análisis Multivariable para las Ciencias Sociales*. Ed. Prentice Hall. Madrid.

López, M. y J. López (1996). *Estadística para actuarios*. Editorial Mapfre, Madrid.

Manual IBM SPSS Regression 2.0.

Peña, D. (2002). *Análisis de datos multivariantes*. Ed. McGraw Hill. Madrid.

Real Decreto Ley 1514/2007 de 16 de Noviembre. Plan General de Contabilidad.

Real Decreto Ley 1515/2007 de 16 de Noviembre. Plan General de Contabilidad de pequeñas y medianas empresas.

DISCUSIÓN DE RESULTADOS

A continuación se analizarán de forma pormenorizada los resultados obtenidos para predecir la insolvencia en los modelos globales y en los modelos sectoriales, utilizando información correspondiente a 1 y 2 años antes de la quiebra.

En relación con los modelos estimados con información disponible 1 año antes de la quiebra (M.I), los modelos obtenidos, tanto los sectoriales como los modelos globales (sin y con dummies) han alcanzado buenos resultados, en términos del porcentaje de clasificación obtenido en la muestra de testeo. Estos porcentajes oscilaron entre el 75,6% logrado por el modelo del sector agricultura, pasando por el 89,2% obtenido por el modelo del sector industrial y llegando al 91,2% registrado por el modelo del sector hostelería.

Hay que apuntar que todos los modelos sectoriales han incorporado variables explicativas que la literatura previa identificó como particulares de ese sector de actividad. Así, en el sector agricultura destacó la variable VMSA2 (EBIT/Gastos Financieros), en el sector industria destacó VMSI7 (Resultado de Explotación/Patrimonio Neto), en el sector construcción VMSC1 (Gastos Financieros/EBIT), en el sector comercio y servicios VMSCS4 (Ventas/Existencias) y en el sector de hostelería VMSH1 y VMSH3 (EBITDA/Pasivo corriente y Total Deuda Financiera/EBITDA).

Por su parte, los modelos globales, sin variables cualitativas sectoriales y con ellas, consiguieron unos

porcentajes de clasificación del 80,8 y del 81,7% en la muestra de testeo, respectivamente.

Al realizar los tests no paramétricos en el análisis exploratorio se comprobó que un gran número de las 34 variables consideradas se mostraban como variables estadísticamente relevantes en la predicción de la insolvencia en los diferentes sectores económicos considerados, siendo pocas las que no han mostrado ningún signo de relevancia. Así, en el sector de agricultura solo con VMG5⁹ y VMG8¹⁰ se aceptó la hipótesis nula de igualdad de parámetros en las distribuciones de empresas solventes e insolventes. En industria, sólo VMG5. En el sector de construcción, VMG3¹¹ y VMG5. En comercio y servicios, VMG3 y VMSCS4¹² y en el sector hostelería, solo VMG9¹³. Finalmente, en la muestra global, únicamente la variable VMG8 aceptó la hipótesis nula en los tests no paramétricos.

Los resultados obtenidos en el análisis confirmatorio confirmaron los obtenidos en el análisis exploratorio, resultando explicativas de los modelos Logit estimados aquellas variables que previamente se habían mostrado como relevantes estadísticamente, con las únicas excepciones VMG5 en el sector industria, y VMG3 y VMSCS4 en el sector comercio y servicios, que resultaron

⁹VMG5=Total Ingresos/Total Activo

¹⁰VMG8= Activo Corriente/Total Activo

¹¹VMG3=Fondo de Maniobra/Total Activo

¹²VMSCS4= Ventas/Existencias

¹³VMG9= Beneficio después de Impuestos/Patrimonio Neto

ser predictores significativos en los modelos sectoriales respectivos de predicción y, sin embargo, no presentaron la suficiente relevancia estadística en los tests no paramétricos.

Así, las variables explicativas de la insolvencia en el sector de agricultura 1 año antes de la quiebra fueron VMG1, VMG6 y VMSA2. En el sector de industria resultaron como variables explicativas VMG1, VMG2, VMG3, VMG4, VMG5, VMG6, VMG8 y VMSI7. En el sector de la construcción resultaron ser las variables VMG1, VMG6, VMG7, VMG8 y VMSC1. En las empresas pertenecientes al sector de comercio y servicios VMG1, VMG2, VMG3, VMG7, VMG8 y VMSCS4. Por su parte, en el sector de la hostelería, los predictores significativos de la insolvencia en M.I fueron VMG1, VMG2, VMG5, VMSH1 y VMSH3. Finalmente, en el modelo global sin dummies, VMG1, VMG3, VMG4 y VMG5 resultaron ser los predictores significativos, y en el modelo global con dummies VMG1, VMG2, VMG3, VMG4, VMG6 y las dummies de comercio y servicios.

Como puede comprobarse, la rentabilidad sobre activos (ROA), con la nomenclatura VMG1, resultó explicativa en todos los modelos estimados. Como era de esperar, en todos los modelos aparece con signo negativo respecto de la insolvencia señalando que, cuanto mayor sea el beneficio neto obtenido por la empresa por unidad monetaria invertida en su activo, menor será la probabilidad de que sea insolvente.

Si analizamos los coeficientes y los Odds ratios obtenidos por estas variables en los distintos modelos estimados se comprueba que es la variable que mayor coeficientes presenta en valores absolutos. Sin embargo, es la que menores Odds ratio registra (0,000 en todos los casos).

Otras tres variables han aparecido como comunes en casi todos los modelos: VMG2, VMG3 y VMG6. En el caso de VMG2 (Activo Corriente/Pasivo Corriente) y VMG6 (Quick Ratio), aparecen como variables altamente correlacionadas en prácticamente todos los sectores económicos, con valores del coeficiente de Pearson superiores a 0,9. Son variables expresivas de la liquidez a corto plazo, lo que viene a evidenciar la importancia de este parámetro en relación con la insolvencia, con independencia del tipo de sector económico del que se trate. En este sentido, resulta oportuno pensar que una de las causas más inmediatas de la insolvencia son los problemas de liquidez, que terminan precipitando el fallo empresarial, entendido como la imposibilidad de atender las obligaciones contraídas con los acreedores. Por su parte, la variable VMG3 incidiría en esta cuestión desde otra perspectiva, al ponerse de manifiesto que una diferencia positiva y creciente de activos corrientes frente a pasivos corrientes, en relación con el tamaño de la empresa, termina teniendo repercusión sobre la insolvencia empresarial.

Parece razonable, pues, que estas tres variables relacionadas con la capacidad de la empresa de atender

sus obligaciones a corto plazo sean las que hayan aparecido prácticamente en todos los modelos y sectores, con información de 1 año antes de la quiebra.

En este mismo sentido, entendemos que las variables representativas del endeudamiento total, de cargas financieras y de recursos generados, como VMG7 (Total Deuda/Total Activo), VMSC1 (Gastos Financieros/EBIT) y VMSH3 (Total Deuda Financiera/EBITDA) hayan aparecido de manera casi exclusiva como predictores significativos de la insolvencia en tres sectores económicos, como en construcción, en comercio y servicios, y en hostelería.

Si comparamos los resultados obtenidos en la literatura previa en los denominados modelos descentrados, o elaborados a partir de muestras heterogéneas de sectores, con los estimados en esta tesis, puede constatar que los resultados obtenidos se sitúan en un rango intermedio, existiendo trabajos previos que muestran mejores y peores resultados. Así, encontramos trabajos que presentan excelentes resultados en la muestra de testeo como el de Shuk-Wern et al. (2011), con un porcentaje de acierto del 90,0%, y otros que se sitúan por debajo de nuestros resultados como el de Charalambous et al. (2000) con un 77,9% en el testeo, e incluso el resultado de Chen et al. (2011) con un porcentaje de acierto en la muestra de testeo del 64,5%.

Con respecto a las variables utilizadas, hay mucha heterogeneidad, sin encontrarse un patrón común entre los

trabajos previos y los modelos globales desarrollados en el presente trabajo de investigación. Hay algunas variables comunes como, por ejemplo, VMG2, VMG4 y VMG5, que han sido utilizadas en el trabajo de Shuk-Wern et al. (2011). También encontramos que VMG1 y VMG3 fueron empleadas también en el trabajo de Chen et al. (2011), no apareciendo ninguna variable en común con el trabajo de Charalambous et al. (2000).

En relación con el modelo estimado para el sector agricultura, tan solo el trabajo de Vavrina et al. (2013) empleó la regresión logística en su estudio, con un porcentaje de clasificación en la muestra de entrenamiento del 71,9% para año antes de la quiebra. En este sentido, nuestro modelo ofrece un mejor resultado, alcanzando el 78,5% en la muestra de entrenamiento y el 75,6% en la muestra de testeo.

Nuestro modelo estimado del sector industrial para M.I. obtuvo un resultado en la muestra de validación del 89,2%, resultado muy similar al obtenido por Lin (2009) con un 89,4% en la muestra de testeo. Tan solo supera nuestro resultado el modelo de Zhang et al. (2013), con un poder de clasificación en la validación del 95,2%. Por debajo de estos resultados se encuentran los trabajos de Zhang et al. (1999), Darayseh et al. (2003) y el de Yi-Chung et al. (2005).

Con el trabajo de Lin (2009) únicamente encontramos en común la variable VMG5. Con el trabajo de Zhang et al. (2013) tan solo se comparten las variables

VMG1 y VMSI7, siendo esta última una de las variables que emergió como específica en los modelos revisados del sector industrial.

En el sector de la construcción nuestro resultado se ha situado en el 81,5% de acierto en la clasificación con la muestra de testeo para M.I., que mejora el resultado obtenido por Mínguez (2006), pues alcanzó sólo un 76,9%. Entre las variables utilizadas, compartimos con Mínguez (2006) la variable VMG1 y con Treewichayapong et al. (2011) las variables VMG1, VMG7 y VMG8.

Por lo que se refiere al sector comercio y servicios, nuestro modelo ha obtenido un porcentaje de clasificación del 83,5% en la muestra de testeo. Este porcentaje es inferior al obtenido por Youn y Zheng (2010), que obtuvieron un 95% de acierto en la muestra de testeo 1 año antes del fallo empresarial. Por debajo de estos resultados se encuentra el modelo de Kim (2011) que obtuvo un porcentaje con la muestra de entrenamiento del 80,0%.

El mejor modelo estimado para M.I. resultó ser el del sector de hostelería, con un porcentaje de acierto en la clasificación del 91,2%. Si bien este es un resultado notable, se encuentra algo por debajo del obtenido por Kim y Gu (2006b) que consiguió un 93%. Dos años antes de la quiebra, nuestro modelo del sector de hostelería registró un porcentaje de acierto del 81,0% en la muestra de testeo. Los trabajos de Kim y Gu (2006a) y Youn y Gu (2010a) registraron, dos años antes de la quiebra, un 84%

y un 85% de acierto con la muestra de entrenamiento. Sin embargo ellos no validaron sus modelos con muestras de testeo.

Cabe destacar que el objetivo de obtener unas conclusiones comparables con los trabajos realizados por el resto de autores sobre predicción de insolvencia nos ha llevado también a utilizar un gran número de variables, que se corresponderían con todas las que inicialmente fueron consideradas en los diferentes trabajos de investigación, inclusive aquéllas que finalmente no resultaron explicativas en los mismos. Este hecho hace que muchas de estas variables sean bastante similares y muestren información con escasas diferencias, presentando altos niveles de correlación entre ellas, si bien no se ha prescindido de ninguna con objeto de obtener unas conclusiones más robustas.

Los resultados obtenidos por los modelos estimados con información disponible de dos años antes de la quiebra obtuvieron peores resultados que los obtenidos para M.I., hecho este habitual en los modelos de predicción de la insolvencia. En concreto, los resultados oscilaron entre el 69,6% de clasificación registrado en el modelo del sector agricultura, el 78,1% obtenido por el modelo del sector industrial, hasta el 81,0% registrado por el modelo del sector hostelería. A diferencia de lo ocurrido en M.I., no todos los modelos estimados en M.II. mostraron variables específicas de sector. Así, con la excepción de los modelos del sector de comercio y servicios y del sector de hostelería, el resto de modelos sectoriales sí contaron con

predictores significativos específicos de su sector. En el sector agricultura destacó la variable VMSA1 (Fondos Propios/Total Deuda), en el sector industria destacó VMSI4 (Fondos Propios/Pasivo No Corriente) y en el sector construcción VMSC1 (Gastos Financieros/EBIT). Por su parte, los modelos globales, sin variables cualitativas sectoriales y con ellas, consiguieron unos porcentajes de clasificación del 71,1 y del 72,7%, respectivamente.

El análisis exploratorio llevado a cabo en M.II. mostró que prácticamente todas las variables consideradas fueron relevantes en la predicción de la insolvencia en los diferentes sectores, siendo muy pocas las que no mostraron ningún signo de relevancia. Así, en el sector de construcción solo para VMG3 y VMG8¹⁴ se aceptó la hipótesis nula de igualdad de parámetros en las distribuciones de empresas solventes e insolventes. En industria, solo para VMSI3¹⁵. En el sector de construcción, para VMG3¹⁶ y VMG5. En comercio y servicios, para VMG9, VMG5 y VMSCS4¹⁷ y VMSCS5¹⁸ y, por último, en el sector hostelería, para VMG3¹⁹. Además, en la muestra global todas las variables rechazaron la hipótesis nula en los tests no paramétricos.

¹⁴VMG8 = Activo Corriente/Total Activo

¹⁵(Activo corriente - Pasivo corriente)/Total Capital

¹⁶VMG3 = Fondo de Maniobra/Total Activo

¹⁷VMSCS4 = Ventas/Existencias

¹⁸ Ventas/Total Activo

¹⁹Fondo de Maniobra/Total Activo

Los resultados obtenidos en el análisis confirmatorio corroboraron los obtenidos en el análisis exploratorio, resultando explicativas en los modelos Logit estimados aquellas variables que previamente se habían mostrado como relevantes estadísticamente, con la única excepción de VMG3 en el sector de la construcción, que resultó ser significativa en el modelo sectorial y, sin embargo, no presentó la suficiente relevancia estadística en los tests no paramétricos.

Así, las variables explicativas de la insolvencia en el sector de agricultura fueron las variables VMG1 y VMSA1, para dos años antes de la quiebra. En el sector de industria resultaron como variables explicativas VMG1, VMG2, VMG3, VMG4, VMG6, VMG7, VMG8, VMG9 y VMSI4. En el sector de la construcción resultaron ser las variables VMG1, VMG3, VMG6, VMG7, VMG8 y VMSC1. En las empresas pertenecientes al sector de comercio y servicios VMG1, VMG2, VMG3 y VMG7. Por su parte, en el sector de la hostelería, los predictores significativos de la insolvencia en M.II. fueron VMG2, VMG4 y VMG6. Finalmente, en el modelo global sin dummies, VMG1, VMG2, VMG3 y VMG9 y en el modelo global con dummies VMG1, VMG2, VMG3, VMG9, y la dummy de industria.

A excepción del modelo del sector de hostelería estimado para M.II., en todos los modelos la rentabilidad sobre activos (ROA), con la nomenclatura VMG1, resultó explicativa. Como era de esperar, en todos aparece con signo negativo respecto de la insolvencia señalando que, cuanto mayor sea el beneficio neto obtenido por la

empresa por unidad monetaria invertida en su activo, menor será la probabilidad de que sea insolvente.

Otras tres variables han aparecido como comunes en casi todos los modelos: VMG2, VMG6 y VMG7. En el caso de VMG2 (Activo Corriente/Pasivo Corriente) y VMG6 (Quick ratio), al igual que en M.I., aparecen como variables altamente correlacionadas en prácticamente todos los sectores económicos, con valores del coeficiente de Pearson superiores a 0,9. Son variables representativas de la liquidez a corto plazo, lo que viene a evidenciar la importancia de este parámetro en relación con la insolvencia, con independencia del tipo de sector económico del que se trate y, lo que resulta más esclarecedor, a un plazo menos inmediato con respecto al desenlace del fallo empresarial, entendido como la imposibilidad de atender las obligaciones contraídas con los acreedores no ya a corto plazo, como en M.I., sino a un mayor plazo como así nos lo ponen de manifiesto los modelos estimados en M.II.

En este plazo de M.II. sí emerge, como una variable destacada en tres de los modelos, la variable VMG7 (Total Deuda/Total Activo), expresiva del efecto que el volumen de endeudamiento genera en el balance de una empresa y su relación con la insolvencia.

Si comparamos los resultados obtenidos en la literatura previa sobre los denominados modelos descentrados, o elaborados a partir de muestras heterogéneas de sectores, con los estimados en esta tesis,

puede constatar que los resultados obtenidos se sitúan en un rango intermedio, existiendo trabajos previos que muestran mejores y peores resultados, siempre y cuando empleen regresión logística. Dos años antes de la quiebra, Charalambous et al. (2000) da un porcentaje de acierto en la muestra de testeo del 68,6%, situándose unos puntos porcentuales por debajo de nuestro modelo.

En relación con el modelo estimado para el sector agricultura, los resultados de nuestro modelo en M.II. muestran un resultado de clasificación con la muestra de testeo del 69,6%, algo mejor que el modelo de Vavrina et al. (2013), que obtiene un 64,4%. Por otra parte, la única variable significativa que aparece en nuestro modelo y que se utiliza también en el trabajo de Vavrina et al. (2013) es VMG1.

El modelo industrial estimado con la muestra M.II registró un resultado del 78,1% en la muestra de testeo. No podemos comparar este resultado ya que sólo Lin (2009) ofrece un porcentaje de clasificación, y es del 71,7% para tres años antes de la quiebra.

En el sector de la construcción, y en referencia a las variables utilizadas, compartimos con Mínguez (2006) la variable VMG1 y con Treewichayapong et al. (2011) las variables VMG1, VMG7 y VMG8.

Por lo que se refiere al sector comercio y servicios para M.II, nuestro modelo clasificó correctamente un 79,5% de las empresas de la muestra, situándose por

debajo del modelo de Youn y Zheng (2010) que, en ese plazo, registró un poder de clasificación del 85,0%. Las únicas variables comunes encontradas con estos estudios son las VMG1, VMG2 y VMG8, que se usaron también en el estudio de Youn y Zheng (2010).

El mejor modelo estimado para M.II. volvió a ser el del sector de hostelería, con un porcentaje de clasificación correcto del 81,0% en la muestra de testeo. Los trabajos de Kim y Gu (2006a) y Youn y Gu (2010a) registraron, dos años antes de la quiebra, un 84% y un 85% de acierto con la muestra de entrenamiento, pero no testaron sus modelos. En M.II., nuestro modelo no incorporó, como variables significativas, ninguna variable específica de sector.

Como consecuencia de las conclusiones anteriores y de la investigación documental y empírica desarrollada creemos que el presente trabajo de investigación contribuye al conocimiento financiero empresarial en diferentes aspectos. En primer lugar, consigue dilucidar una cuestión que, aunque ya apuntada por otros autores no había sido objeto de estudio específico con un diseño, un método y una muestra específica.

En segundo lugar, permite sacar conclusiones a la hora de abordar una estrategia de predicción de la insolvencia en diferentes sectores económicos. Prácticamente en todos los casos, un modelo global es capaz de predecir con éxito la insolvencia en muestras de empresas pertenecientes a sectores económicos concretos,

lo que implicaría un ahorro de costes considerable en la elaboración y desarrollo de numerosos modelos sectoriales.

En tercer lugar, este trabajo ha puesto de manifiesto el concurso de variables explicativas sectoriales, de carácter específico, que permiten tener un conocimiento mucho más concreto de los factores que explican la insolvencia. El conocimiento de estas variables como específicamente sectoriales en la explicación del fracaso empresarial puede ayudar a los agentes económicos y usuarios de esta información a actuar preventivamente.

Por último, y dado que el primer paso en la gestión de los riesgos es realizar una medición de los mismos, una adecuada puntuación del riesgo de quiebra puede ayudar en este sentido. Por ello, antes de la decisión de utilizar un determinado modelo, la empresa necesita fundamentos que le orienten en las limitaciones de éstos y sobre cuál de ellos (globales o sectoriales) se adapta mejor a sus circunstancias.

CONCLUSIONES

El objetivo de la presente tesis doctoral es cubrir el gap existente en la literatura en relación con la cuestión de la supuesta superioridad de los modelos descentrados, frente a los centrados, a la hora de predecir la insolvencia empresarial. Se ha tratado de dilucidar esta cuestión con un diseño ad-hoc, superando la ausencia de conclusiones definitivas sobre esta cuestión en la literatura previa debido a la disparidad de métodos, enfoques, bases de datos disponibles, periodos temporales y países.

Para resolver esta cuestión, se ha tomado una base de datos, formada tanto por empresas solventes como insolventes, pertenecientes a cinco sectores diferenciados: agricultura, industria, construcción, comercio y servicios, y sector de hostelería, a partir de los datos proporcionados por SABI (Sistema de Análisis de Balances Ibéricos). Los datos utilizados para este estudio se corresponden al periodo comprendido entre el año 2010 y el año 2012.

Se han construido modelos sectoriales o centrados y modelos globales o descentrados para un año y dos años antes de la entrada en quiebra (M.I y M.II). Se han considerado un total de 1.287 y 1.290 empresas en situación de quiebra, para M.I y M.II, respectivamente.

La metodología aplicada para la obtención de los modelos ha sido la construcción de modelos de regresión logística, y para la estimación de dichos modelos se emplearon 34 variables financieras, así como variables dummies representativas de los sectores de actividad bajo consideración. Con esta metodología se obtuvo un poder

de clasificación que ha oscilado entre un 69,6% (Sector Agricultura en M.II) y un 91,2% (Sector Hostelería en M.I).

Para la muestra de empresas 1 año antes de la entrada en quiebra (M.I.), se estimaron dos modelos globales o descentrados. Uno sin la inclusión de variables cualitativas sectoriales y otro con la consideración de variables dummies sectoriales.

Por lo que se refiere al modelo global sin dummies, las variables que resultaron ser significativas en el mismo fueron VMG1, VMG3, VMG4 y VMG5. El modelo presentó un poder de clasificación en la muestra de testeo del 80,8%.

El segundo modelo global estimado, con la inclusión de variables dummies sectoriales, presentó como variables significativas VMG1, VMG2, VMG3, VMG4, VMG6, dummy del sector industria y dummy del sector comercio y servicios. El porcentaje de clasificación que obtuvo este modelo en la muestra de testeo alcanzó el 81,7%, con lo que es posible inferir que la inclusión de variables cualitativas sectoriales mejora la capacidad predictiva del modelo global o descentrado.

Como se ha mencionado, se estimaron igualmente cinco modelos sectoriales. El modelo de agricultura para M.I contenía las variables VMG1, VMG6 y VMS2, alcanzando un resultado de clasificación en la muestra de testeo del 75,6%. El modelo de industria ha sido el que

más variables significativas ha empleado (VMG1, VMG2, VMG3, VMG4, VMG5, VMG6, VMG8 y VMSI7). De la misma forma, ha sido el segundo modelo que mejores resultados de clasificación ha conseguido, con un 89,2% en la muestra de testeo. Por su parte, el modelo del sector de la construcción ha estado formado por las variables VMG1, VMG6, VMG7, VMG8 y VMSC1, obteniendo un porcentaje de clasificación del 81,5%. El modelo estimado para el sector de comercio y servicios ha mostrado como significativas las variables VMG1, VMG2, VMG3, VMG7, VMG8 y VMSCS4, con un poder de predicción del 83,5%. Finalmente, el modelo estimado para el sector de hostelería empleó las siguientes variables significativas: VMG1, VMG2, VMG5, VMSH1 y VMSH3, registrando el porcentaje de clasificación más alto de todos los modelos sectoriales (91,2% en la muestra de testeo). Hay que hacer notar, además, que todos los modelos sectoriales contaron en sus modelos con variables específicas sectoriales que resultaron ser significativas. Asimismo, la variable VMG1 aparece en todos los modelos estimados.

Igualmente, para la muestra de empresas dos años antes de la entrada en quiebra (M.II), se estimaron dos modelos globales o descentrados. Por lo que se refiere al modelo global sin dummies, las variables que resultaron ser significativas en el mismo fueron VMG1, VMG2, VMG3 y VMG9. El modelo presentó un poder de clasificación en la muestra de testeo del 71,1%.

El segundo modelo global estimado, con la inclusión de variables dummies sectoriales, presentó como variables

significativas VMG1, VMG2, VMG3, VMG9 y la dummy del sector industria. El porcentaje de clasificación que obtuvo este modelo en la muestra de testeo alcanzó el 72,7%, con lo que es posible inferir, al igual que ocurrió en M.I, que la inclusión de variables cualitativas sectoriales mejora la capacidad predictiva del modelo global o descentrado.

Se estimaron igualmente cinco modelos sectoriales para M.II. El modelo de agricultura con las variables VMG1 y VMSA1 ha alcanzando un resultado de clasificación en la muestra de testeo del 69,6%. De la misma forma que en M.I, el modelo de industria ha sido el que más variables significativas ha empleado (VMG1, VMG2, VMG3, VMG4, VMG6, VMG7, VMG8, VMG9 y VMSI4). De la misma forma, ha sido el tercer modelo que mejores resultados de clasificación ha conseguido, con un 78,1% en la muestra de testeo. El modelo del sector de la construcción ha incluido seis variables (VMG1, VMG3, VMG6, VMG7, VMG8 y VMSC1), obteniendo un porcentaje de clasificación del 74,5%. Por su parte, el modelo estimado para el sector de comercio y servicios contenía cuatro variables significativas (VMG1, VMG2, VMG3 y VMG7), arrojando un poder de clasificación del 79,5%. Finalmente, el modelo estimado para el sector de hostelería empleó las variables VMG2, VMG4 y VMG6, registrando, al igual que en M.I, el porcentaje de clasificación más alto de todos los modelos sectoriales (81,0% en la muestra de testeo).

Hay que hacer notar que en M.II no todos los modelos contaron, entre las significativas, con variables específicas sectoriales. Concretamente, los modelos del

sector de comercio y servicios y del sector hostelería no presentaron, en esta ocasión, variables significativas específicas de su sector.

Conforme a lo anterior, los resultados empíricos obtenidos nos han permitido aceptar la hipótesis H_1 , pues se han confirmado la existencia de diferencias entre los modelos globales y sectoriales (el conjunto de regresores que contienen son diferentes, aun siendo algunos comunes). Esta hipótesis ha sido verificada tanto para M.I como para M.II.

Asimismo, y en relación a la hipótesis H_2 , la evidencia empírica, tanto para M.I como para M.II, sugiere que la inclusión de variables cualitativas sectoriales mejora los modelos globales estimados, a tenor de la comparación efectuada mediante los criterios de Akaike, Schwartz y Hannann-Quinn. Por tanto, se confirma la segunda hipótesis planteada (H_2).

Y en relación con la hipótesis H_3 (el modelo global es capaz de predecir correctamente con las muestras de los diferentes sectores), la evidencia empírica aconseja aceptarla solo parcialmente, pues el modelo del sector industria en M.II es superior al modelo global testeado en la muestra del sector industrial. Para el resto de modelos, los modelos globales estimados en M.I y M.II son capaces de obtener buenos resultados de predicción en las muestras sectoriales.

Finalmente, y como consecuencia de los resultados del presente trabajo, surgen futuras líneas de investigación que consideramos de suma utilidad para perfeccionar los modelos de predicción de insolvencia. Así, y en primer lugar, sería interesante comprobar si las conclusiones obtenidas para la muestra de empresas españolas son verificadas en empresas de otros ámbitos geográficos, lo que dotaría a estas conclusiones de una elevada capacidad de generalización.

Del mismo modo, también podría resultar relevante modificar la definición de los sectores considerados, y de esta forma comprobar si las conclusiones aquí obtenidas serían las mismas en sectores económicos más o menos ampliamente definidos.

Además, y dado que los modelos globales contruidos son capaces de predecir la quiebra de una compañía con exactitud, podría investigarse sobre la utilidad de los mismos para ayudar a los auditores externos a evaluar la viabilidad futura de las empresas (going concern).

Por último, y como la eficacia de los modelos posiblemente variará en función de las condiciones macroeconómicas, sería interesante conocer los resultados de clasificación de dichos modelos en diferentes etapas del ciclo económico.

BIBLIOGRAFÍA GENERAL

Abdullah, N.A.H., Halim, A., Ahmad, H. y R. M. Rus (2008). Predicting Corporate Failure of Malaysia's listed companies: comparing multiple discriminant analysis, logistic regression and the hazard model. *International Research Journal of Finance and Economics*, 15, 201-217.

Abidali, A.F. y F. Harris. (1995). A methodology for predicting company failure in the construction industry. *Construction Management and Economics*, 13, 189-196.

Aharony, J., Jones, Ch. P. y I. Swary. (1980). An analysis of risk and return characteristics of Corporate Bankruptcy using capital market data. *The Journal of Finance*, 35, 4, 1001-1016.

Ahn H. y K. Kim. (2009). Bankruptcy prediction modelling with hybrid case-based reasoning and genetic algorithms approach. *Application Software Computing*, 9, 599-607.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov and F. Csaki (Eds.), *Second international symposium on information theory*, 267-281.

Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, 19, 6, 716-723.

Akers, M. (2007). A Review of Bankruptcy Prediction Studies: 1930 to Present. *Journal of Financial Education*, 33 (Winter), 1-42.

Alam, P., Booth, D., Lee, K. y T. Thordarson (2000). The use of fuzzy clustering algorithm and self-organizing neural networks for identifying potentially failing banks: An experimental study. *Expert Systems with Applications*, 18, 185-199.

Alici, P.Y. (1996). *Neural networks in corporate failure prediction. The UK experience*. Neural Networks in Financial Engineering. Mostafa, A. Moody, J. y Weigend, A. (eds). Singapore. *World Scientific*, 393-406.

Allen, D.M. (1974). The relationship between variable selection and data augmentation and a method for prediction. *Technometrics*, 16, 125-127.

Altman, E.I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23, 4, 589-609.

Altman, D.G. (1991). *Practical Statistics for Medical Research*. Londres. Chapman.

Altman, E.I. (1993). *Corporate Financial Distress and Bankruptcy*. 2nd ed. New York. John Wiley & Sons.

Altman E.I., Marco, G. y F. Varetto (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (The Italian experience). *Journal of Banking and Finance*, 18, 505-529.

Altman, E.I., Y. H. Eom y D.W. Kim (1995). Failure Prediction: Evidence from Korea. *Journal of International Financial Management and Accounting*, 6, 230-249.

Altman, E.I. y E. Hotchkiss. (2006). *Corporate Financial Distress and Bankruptcy*. Predict and avoid Bankruptcy, Analyze and Invest in Distressed Debt. Third Edition. Wiley Finance.

Aggarwal, R. y J.W. Goodell (2014). Cross-national differences in access to finance: Influence of culture and institutional environments. *Research in International Business and Finance*, 31, 193–211.

Anandarajan, M., Lee, P. y A. Anandarajan (2001). Bankruptcy prediction of financially stressed firms: An examination of the predictive accuracy of artificial neural networks. *Intelligent Systems Accounting Finance Management*, 10, 69–81.

Appetiti, A. (1984). Identifying unsound firms in Italy. An attempt to use trend variables. *Journal of Banking and Finance*, 8, 2, 269-279.

Appiah, K.O. y J. Abor. (2009). Predicting corporate failure: some empirical evidence from the UK. *Benchmarking: An International Journal*, 16, 3, 432-444.

Arkaradejdachachai, C. (1993). Study of corporate turnaround. Using a probability of bankruptcy. *Tesis doctoral*. Universidad de Missouri.

Arquero, J.L., Abad, M.C. y S.M. Jiménez (2009). Procesos de fracaso empresarial en Pymes. Identificación y contrastación empírica. *Revista Internacional de la Pequeña y Mediana Empresa*, 2, 1, 64-77.

Atiya, A.F. (2001). Bankruptcy Prediction for Credit Risk Using Neural Network: A Survey and New Results. *IEEE Transactions on Neural Networks*, 12, 4, 929-935.

Aziz, A., Emanuel, D.C. y G.H. Lawson (1988). Bankruptcy Prediction. An investigation of cash flow based models. *Journal of Management Studies*, 25, 5, 419-437.

Aziz, A. y G. Lawson (1989). Cash Flow Reporting and Financial Distress Models: Testing of Hypotheses. *Financial Management*, 18, 1, 55-63.

Back, B., Laitinen, T. y K. Sere (1996). Neural networks and genetic algorithms for bankruptcy predictions. *Expert Systems with Applications*, 11, 407-413.

Baldwin, J. y G.W. Glezen (1992). Bankruptcy prediction using quarterly financial statement data. *Journal of Accounting, Auditing and Finance*, 7, 269-289.

Bartoloni, E. y M. Baussola (2014). Financial Performance in Manufacturing Firms: A comparison between parametric and non-parametric approaches, *Business Economics*, 49, 1, 32-45.

Beaver, W.H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 5, 71-111.

Bell, T., Ribar, G. y J. Verchio (1990). Neural nets versus logistic regression: A comparison of each model's ability to predict commercial bank failures. Proceedings of the 1990 D&T, University of Kansas Symposium on Auditing Problems.

Bellovary, J., Giacomino, D. y M. Akers (2007). A Review of Bankruptcy Prediction Studies: 1930 to Present. *Journal of Financial Education*, 33 (Winter), 1-42.

Bernstein, L.A. (1984). *Análisis de Estados Financieros*. Ed. Deusto. Bilbao.

Betz, F., Oprică, S., Peltonen, T.A. y P. Sarlin (2014). Predicting distress in European banks. *Journal of Banking and Finance*, 45, 225–241.

Bhargava, M., Dubelaar, C. y T. Scott (1998). Predicting bankruptcy in the retail sector: an examination of the validity of key measures of performance. *Journal of Retailing and Consumer Services*, 5, 2, 105-117.

Bolton, R. y D. Hand (2002). Statistical fraud detection: a review. *Statistical Science*, 17(3), 235-255.

Bongini, P., Ferri, G. y H. Hahm (2000). Corporate bankruptcy in Korea. Only the strong survive?. *The Financial Review*, 35, 4, 31–50.

Boritz, J.E., Kennedy, D.B. y A. Albuquerque (1995). Predicting corporate failure using a neural network approach. *Intelligent Systems in Accounting, Finance and*

Management, 4, 95–111.

Boritz, J. y D. Kennedy (1995). Effectiveness of Neural Network Types for Prediction of Business Failure. *Expert Systems with Applications*, 9, 4, 503-512.

Borrajo, M., Baruque, B., Corchado, E., Bajo, J. y J. Corchado (2011). Hybrid neural intelligent system to predict business failure in small-to-medium-size enterprises. *International Journal Neural Systems*, 21, 4, 277–296.

Bose, I. y R. Pal (2006). Predicting the survival or failure of click-and-mortar corporations: a knowledge discovery approach. *European Journal of Operations Research*, 174, 959–982.

Brigham, E.F. y L.C. Gapenski (1994). *Financial Management: Theory and Practice*. Orlando, FL: The Dryden Press.

Callejón, A.M., Casado, A.M., Fernández, M.A. y J.I. Peláez (2013). A System of Insolvency Prediction for industrial companies using a financial alternative model with neural networks. *International Journal of Computational Intelligence Systems*, 4, 1-13.

Camillo, A.A., Connolly D.J. y W.G. Kim (2008). Success and Failure in Northern California. Critical success factors for independent restaurants. *Cornell Hospitality Quarterly*, 49, 364-380.

Casey, C.J. Jr. (1980). The usefulness of accounting ratios for subjects' predictions of corporate failure: Replication and Extensions. *Journal of Accounting Research*, 18, 2, 603-613.

Casey, C. y N. Bartczak (1985). Using Operating Cash Flow Data to Predict Financial Distress: Some Extensions. *Journal of Accounting Research*, 23, 1, 384-401.

Celentani, M., García-Posada, M. y F. Gómez (2011). *The Spanish Business Bankruptcy Puzzle and the Crisis*. Fundación de Estudios de Economía Aplicada. Documento de Trabajo, nº 2010-2011.

Chang Lee, K., Han, I. y Y. Kwon (1996). Hybrid neural network models for bankruptcy predictions. *Decision Support Systems*, 18, 63-72.

Charalambous, C., Chatitou, A. y F. Kaourou (2000). Comparative analysis of artificial neural network models: Application in bankruptcy prediction. *Annals of Operations Research*, 99, 1, 403-419.

Charitou, A., Neophytou, E. y C. Charalambous (2004). Predicting corporate failure. Empirical evidence for the UK. *European Accounting Review* 13, 3, 465–497.

Chauhan, N., Ravi, V. y D.K. Chandra (2009). Differential evolution trained wavelet neural networks: Application to bankruptcy prediction in banks. *Expert Systems with Applications*, 36, 4, 7659-7665.

Chava, S. y R.A. Jarrow (2004). Bankruptcy Prediction with Industry Effects. *Review of Finance*, 8, 537-569.

Chen, M.Y. (2011). Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches. *Computers and Mathematics with Applications*, 62, 4514-4524.

Chen, S., Härdle, W.K. y R.A. Moros (2011). Modeling default risk with support vector machines. *Quantitative Finance*, 11, 1, 135-154.

Chen, M.Y. (2011). Predicting corporate financial distress based on integration of decision tree classification and logistic regression. *Expert Systems with Applications*, 38, 11261-1127.

Cheng, C., Chen, C.L. y C.J. Fu (2006). Financial Distress Prediction by a Radial Basis Function Network with Logit Analysis Learning. *Computer and Mathematics with Applications*, 51, 579-588.

Cheng, S. y M. Pecht (2012). Using cross-validation for model parameter selection of sequential probability ratio test. *Expert Systems with Applications*, 39, 8467–8473.

Cho, M. (1994). Predicting business failure in the hospitality industry: An application of logit model. *PhD Dissertation*. Virginia Polytechnic Institute and State University.

Cho, S., Hong, H. y B.Ch. Ha (2010). A hybrid approach based on the combination of variable selection using decision trees and case-based reasoning using the Mahalanobis distance: For bankruptcy prediction. *Expert Systems with Applications*, 4, 3482-3488.

Chye, K., Chin, T. y G. Peng (2004). Credit scoring using data mining techniques. *Singapore Management Review*, 26,2, 25-47

Cielen, A., Ludo, P. y K. Vanhoof (2004). Bankruptcy prediction using a data envelopment analysis. *European Journal of Operations Research*, 154, 526–532.

Coats, P.K., y L.F. Fant (1993). Recognizing financial distress patterns using a neural network tool. *Financial Management*, 22, 3, 142-155.

Coats, P.K. y L.F. Fant (1992). A neural network approach to forecasting financial distress. *The Journal of Business Forecasting Methods and Systems*, 10, 4, 9-12.

D'Antoni, J., Mishra, A. y S. Chintawar (2009). *Predicting Financial Stress in Young and Beginning Farmers in the United States*. Department of Agricultural Economics and Agribusiness, Louisiana State University.

Darayseh, M., Waples, E. y D. Tsoukalas (2003). Corporate failure for manufacturing industries using firms specifics and economic environment whit logit analysis. *Managerial Finance*, 29, 8, 23-37.

Deakin, E.B. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research*, 10, 1, 167-179.

De Andrés, J., Landajo, M. y P. Lorca (2005). Forecasting business profitability by using classification techniques. A comparative analysis based on a Spanish case. *European Journal of Operational Research*, 167, 2, 518–542.

De Andrés, J., Landajo, M. y P. Lorca (2012). Bankruptcy prediction models based on multinorm analysis: An alternative to accounting ratios, *Knowledge-Based Systems*, 30, 67-77.

Diamond, J. H. (1976). Pattern recognition and the detection of corporate failure. *PhD. Dissertation*, New York University.

Dietrich, J., Arcelus, F.J. y G. Srinivisan (2005). Predicting Financial Failure: Some Evidence from New Brunswick Agricultural Co-ops. *Annals of Public and Cooperative Economics*, 76, 2, 179-194.

Dietrich, R. (1984). Discussion of the methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* (Supplement 1984).

Dimitras, A., S. Zanakis y C. Zopounidis (1996). A survey of business failures with an emphasis on prediction methods and industrial applications. *European Journal of Operational Research*, 90, 487-513.

Dimitras, A.I., Slowinski, R. y R. Susmaga (1999). Business failure prediction using rough sets. *European Journal of Operations Research*, 114, 263–280.

Ding, Y., Song, X. y Y. Zeng (2008). Forecasting financial condition of Chinese listed companies based on support vector machine. *Expert Systems with Applications*, 34, 3081–3089.

Edminster, R. (1972). An empirical test of financial ratio analysis for small business failure prediction. *Journal of Financial and Quantitative Analysis*, 7, 2, 1477-1493.

El Hennawy, R. y R. Morris (1983). The significance of base year in developing failure prediction models. *Journal of Business Finance and Accounting*, 10, 2, 209-223.

El-Temtamy, O. (1995). Bankruptcy prediction: A comparative study on logit and neural networks. *PhD Dissertation*. Middle Tennessee University.

Espahbodi, P. (1991). Identification of problem banks and binary choice models. *Journal of Banking and Finance*, 15, 1, 53-71.

Etemadi, H., Rostamy, A. y H. Dehkordi (2009). A genetic programming model for bankruptcy prediction: empirical evidence from Iran. *Expert Systems Applications*, 36, 2, 3199–3207.

Evans, J. y A.L. Borders (2014). Strategically surviving bankruptcy during a global financial crisis: The importance

of understanding chapter 15. *Journal of Business Research*, 67, 1, 2738–2742.

Fernández, M.A.; Gil, A.M. y P. López (2009). *Estructura y Equilibrio financiero de las PYMES Españolas ante las Nuevas Exigencias de Información Financiera*. Fundación de las Cajas de Ahorros. Madrid.

Friedman, H., Altman, E.I. y D. Kao (1985). Introducing recursive partitioning for financial classification: the case of financial distress. *Journal of Finance*, 40, 269–291.

Fletcher, D. y E. Goss (1993). Forecasting with neural networks. An application using bankruptcy data. *Information and Management*, 24, 159-167.

García, V. y M.A. Fernández (1992). *Solvencia y Rentabilidad de la Empresa Española*. Instituto de Estudios Económicos. Madrid.

González, J., Perea, J. C. y A. Y. Sánchez (2003). El Reglamento sobre procedimientos de insolvencia de la Unión Europea. *Partida Doble*, 141, 6-15.

Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of American Statistical Association*, 70, 320–328.

Gentry, J., Newbold, P. y D. Whitford (1985). Predicting Bankruptcy: If Cash Flow´s Not the Bottom Line, What Is?. *Financial Analysts Journal*, 41, 5, 47-56.

Gesel, T., Baesens, B., Suykens, J., Poel, D., Baestaens, D. y M. Willekens (2006). Bayesian kernel based classification for financial distress detection. *European Journal of Operations Research*, 172, 979–1003.

Gepp, A., Kumar, K. y S. Bhattacharya (2010). Business failure prediction using decision trees. *Journal of Forecasting*, 29, 6, 536–555.

Gilbert, L. R., Menon, K. y K.B. Schwartz (1990). Predicting bankruptcy for firms in financial distress. *Journal of Business Finance and Accounting*, 17, 1, 161-192.

Gill de Albornoz, B. y B. Giner (2013). Predicción del fracaso empresarial en los sectores construcción e inmobiliario: modelos generales versus específicos. *Universia Business Review*, Tercer Trimestre, 118-131.

Gombola, M.J., Haskins, M.E., Ketz, J. E. y D.D. Williams (1987). Cash Flow in Bankruptcy Prediction. *Financial Management*, 16, 4, 55-65.

Gómez, M^a Elena, De la Torre, J.M. y I. Román (2008). Análisis de sensibilidad temporal en los modelos de predicción de insolvencia: una aplicación a las pymes industriales, *Revista Española de Financiación y Contabilidad*, Vol. XXXVII, 137, 85-111.

González, J., Perea, J. C. y A. Y. Sánchez (2003). El Reglamento sobre procedimientos de insolvencia de la Unión Europea. *Partida Doble*, 141, 6-15.

Grover, J. (2003). Validation of a cash flow model. A non Bankruptcy approach. *Ph.D.Dissertation*, Nova Southeastern University.

Grüenberg, M. y O. Lukason (2014). Predicting Bankruptcy of Manufacturing Firms. *International Journal of Trade, Economics and Finance*, 5, 1, 93-97.

Gu, Z. (2002). Analyzing bankruptcy in the restaurant industry: A multiple discriminant model. *International Journal of Hospitality Management*, 21, 1, 25-42.

Gu, Z. y L. Gao (2000). A multivariate model for predicting business failures of hospitality firms. *Tourism and Hospitality Research. The Survey Quarterly Review*, 2, 1, 37-49.

Guan, Q. (1993). Development of optimal network structures for back-propagation-trained neural networks. *PhD dissertation*. University of Nebraska.

Habilollah, J. y S. Farnoosh (2009). The comparison Artificial Neural Networks and Multi Decimal Analysis Models for Forecasting Bankruptcy and Financial Distress, *Proceedings of the World Congress on Engineering*, vol. II, London.

Hair, J.F., Anderson R.E., Tatham, R.L. y W.C. Black (1999). *Análisis multivariante*, 5ª edición. Editorial Prentice Hall. Madrid.

Hair, J. (2008). *Análisis Multivariante*. Editorial Prentice Hall.

Hannan, E.J. y B.G. Quinn (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society, Series B*, 41, 2, 190-195.

Hanweck, G. (1977). Predicting bank failures. Research Papers in Banking and Financial Economics, Financial Studies Section, Board of Governors of the Federal Reserve System, Washington D.C.

Hastie, T., Tibshirani, R. y J. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York.

He, Y. y R. Kamath (2006). Business Failure Prediction in Retail Industry: An empirical evaluation of generic bankruptcy models. *Academy of Accounting and Financial Studies Journal*, 10, 2, 97-110.

Henebry, K. (1996). Do cash flow variables improve the predictive accuracy of a Cox proportional hazards model for bank failure?. *The Quarterly Review of Economics and Finance*, 36, 3, 395-409.

Hing-Ling Lau, A. (1987). A Five-State Financial Distress Prediction Model. *Journal of Accounting Research*, 25, 1, 127-138.

Hopwood, W., McKeown, J. y J. Mutchler (1989). A Test of the Incremental Explanatory Power of Opinions Qualified

for Consistency and Uncertainty. *The Accounting Review*, 44, 1, 28-48.

Hu, Y.C. y F.M. Tseng (2005). Applying Back propagation Neural Network to Bankruptcy Prediction. *International Journal of Electronic Business Management*, 3, 2, 79-103.

Hui, X.-F. y J. Sun (2006). An application of support vector machine to companies financial distress prediction. Lecture Notes. *Artificial Intelligent*, 3885, 274–282.

Huo, Y. (2006). Bankruptcy Situation Model in Small Business: the case of restaurant firms, *Hospitality Review*, 24, 2, 49-58.

Izan, H. (1984). Corporate distress in Australia. *Journal of Banking and Finance*, 8, 303-320.

Jae Kwon Bae (2012). Predicting financial distress of the South Korean manufacturing industries, *Expert Systems with Applications*, 39, 9159-9165.

Jo, H. y I. Han (1996). Integration of case-based forecasting, neural network, and discriminant analysis for bankruptcy prediction. *Expert Systems with Applications*, 11, 415–422.

Jo, H. y I. Han (1997). Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis. *Expert Systems with Applications*, 13, 97–108.

Johnsen, T. y R.W. Melicher (1994). Predicting Corporate Bankruptcy and Financial Distress: Information Value Added by Multinomial Logit Models. *Journal of Economics and Business*, 46, 269-286.

Jones, S. y D.A. Hensher (2007). Forecasting Corporate Bankruptcy: Optimizing the Performance of the Mixed Logit Model. *Abacus*, 43, 241-264.

Karels G.V. y Arun J. Prakash (1987). Multivariate Normality and Forecasting of Business Bankruptcy. *Journal of Business Finance and Accounting*, 14, 4, 573-593.

Kahya, E. y P. Theodossiou (1999). Predicting Corporate Financial Distress: A time series CUSUM Methodology. *Review of Quantitative Finance and Accounting*, 13, 323-345.

Keener, M. (2013). Predicting the financial failure of retail companies in the United States, *Journal of Business and Economics Research*, 11, 8, 373-380.

Ketz, J.E. (1978). The effect of general price-level adjustments on the predictive ability of financial ratios. *Journal of Accounting Research*, 16, 273-284.

Kim, H y Z. Gu (2006 a). A logistic regression analysis for predicting bankruptcy in the Hospitality Industry. *The Journal of Hospitality Financial Management*, 14, 1, 17-34.

Kim, H y Z. Gu (2006 b). Predicting Restaurant Bankruptcy. A Logit Model in Comparison with a

Discriminant Model. *Journal of Hospitality and Tourism Research*, 30, 4, 474-493.

Kim, S.Y. (2011). Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression and multivariate analysis. *The Service Industries Journal*, 31, 3, 441-468.

Koh, H. y L.N. Killough (1990). The use of multiple discriminant analysis in the assessment of the going-concern status of an audit client. *Journal of Business Finance and Accounting*, 17, 2, 179-192.

Kolmogorov, A.N. (1957). On the representations of continuous functions of many variables by superpositions of continuous functions of one variable and addition. *Doklady Akademii Nauk USSR*, 114,5, 953-956.

Korol, T. (2013). Early warning models against bankruptcy risk for Central European and Latin American enterprises. *Economic Modelling*, 31, 22-30.

Kotler, P., Bowen, J. y J. Makens (1996). *Marketing for Hospitality & Tourism*. Upper Saddle River, Editorial Prentice Hall.

Kwak, W., Shi, Y. y G. Kou (2011). Bankruptcy prediction for Korean firms after the 1997 financial crisis: using a multiple criteria linear programming data mining approach. *Review of Quantitative Finance and Accounting*, 38, 4, 441-453.

Kwak, W., Yong Shi y K. Gang (2012). Bankruptcy prediction for Korean firms after the 1997 financial crisis: using a multiple criteria linear programming data mining approach. *Review of Quantitative Finance and Accounting*, 38, 441-453.

Lane, W., Looney, S. y J. Wansley (1986). An application of the Cox proportional hazards model to bank failure. *Journal of Banking and Finance*, 10, 511-531.

Langford, D., Iyagba, R. y D.M. Komba (1993). Prediction of solvency in construction companies. *Construction Management and Economics*, 11, 317-325.

Lau, A. (1987). A five-state financial distress prediction model. *Journal of Accounting Research*, 25, 1, 127-138.

Lee, K.C., Han, I. y Y. Kwon (1996). Hybrid Neural Network models for bankruptcy predictions. *Decision Support Systems*, 18, 63-72.

Lee, K. (2001). Pattern classifications and clustering algorithms with supervised and unsupervised neural networks in financial applications. *PhD Dissertation*. Kent State University.

Lee, S., Koh, Y., y K.H. Hang (2011). Moderating effect of capital intensity on the relationship between leverage and financial distress in the US restaurant industry. *International Journal of Hospitality Management*, 30, 429-438.

Leshno, M. y Y. Spector (1996). Neural Network Prediction Analysis: The bankruptcy case. *Neurocomputing*, 10, 125-147.

Levitan, A. y J. Knoblett (1985). Indicators of Exceptions to the Going Concern Assumption. *Auditing: A Journal of Practice and Theory*, 5, 1, 26-39.

Levy, J.P. y J. Varela (2003). *Análisis Multivariable para las Ciencias Sociales*. Ed. Prentice Hall. Madrid.

Ley 22/2003 de 9 de Julio, Concursal.

Ley 38/2011, de 10 de Octubre, de reforma de la Ley 22/2003, de 9 de Julio, Concursal.

Li, H., Huang, H., Sun, J. y C. Lin (2010). On sensitivity of case-based reasoning to optimal feature subsets in business failure prediction. *Expert System with Applications*, 37, 7, 4811–4821.

Libby, R. (1975). Accounting ratios and the prediction of failure: Some behavioral evidence. *Journal of Accounting Research*, 13, 1, 150-161.

Li, H., Sun, J. y J. Wu (2010). Predicting business failure using classification and regression tree: an empirical comparison with popular classical statistical methods and top classification mining methods. *Expert Systems with Applications*, 37, 8, 5895–5904.

Lin, T.H. (2009). A cross model study of corporate financial distress prediction in Taiwan: Multiple Discriminant Analysis, logit, probit and neural networks models. *Neurocomputing*, 72, 3507-3516.

Lin, F., Yeh, C.C. y M.Y. Lee (2011). The use of hybrid manifold learning and support vector machines in the prediction of business failure. *Knowledge-Based Systems*, 24, 95–101.

Lincoln, M. (1984). An empirical study of the usefulness of accounting ratios to describe levels of insolvency risk. *Journal of Banking and Finance*, 8, 321-340.

Lo, A.W. (1986). Logit versus discriminant analysis: a specification test and application to corporate bankruptcies. *Journal of Econometrics*, 31, 151-178.

Logue, L. y J. Merville (1972). Financial policy and market expectations. *Financial Management*, 1, 3, 37-44.

López, M. y J. López (1996). *Estadística para actuarios*. Editorial Mapfre, Madrid.

Lussier, R.N. y S. Pfeiffer (2000). A comparison of business success versus failure variables between U.S. and central eastern Europe Croatian entrepreneurs. *Entrepreneurship Theory and Practice*, 24, 4, 59–67.

Maghyereh, A.I. y B. Awartani (2014). Bank distress prediction: Empirical evidence from the Gulf Cooperation

Council countries. *Research in International Business and Finance*, 30, 126–147.

Malcolm, S. y D. Liou (2007). Industrial sector and financial distress. *Managerial Auditing Journal*, 22, 4, 376-391.

Manual IBM SPSS Statistics. Version 22.

Martin, D. (1977). Early warning of bank failures: A logit regression approach. *Journal of Banking and Finance*, 1, 249-276.

Martin del Brio, B. and C. Serrano-Cinca (1995). Self-organizing neural networks: The financial state of Spanish companies. Article in *Neural Network in the Capital Markets*, Refenes (ed.). Chichester: Wiley, 341-357.

Mateos-Ronco, A., Marín-Sánchez, M^a del Mar, Marí-Vidal, Sergio y E. Seguí-Mas (2011). Los modelos de predicción del fracaso empresarial y su aplicabilidad en cooperativas agrarias. *CIRIEC-España, Revista de Economía Pública, Social y Cooperativa*, 70, abril, 179-208.

McDonald, B. y M.H. Morris (1984). The statistical validity of the ratio method in financial analysis: An empirical examination. *Journal of Business Finance and Accounting*, 11, 1, 89-97.

McGurr, P.T. y DeVaney, S.A. (1998). Predicting business failure of retail firms: An analysis using mixed industry models. *Journal of Business Research*, 43, 169-176.

McKee, T. (1995). Predicting Bankruptcy via Induction. *Journal of Information Technology*, 10, 26-36.

McKee, T.E. y M. Greenstein (2000). Predicting bankruptcy using recursive partitioning and a realistically proportioned data set. *Journal of Forecasting*, 19, 219-230.

McKee, T.E. (2000). Developing a bankruptcy prediction model via rough sets theory. *Intelligent Systems Accounting Finance Management*, 9, 159-173.

Medina, E. (2003). *Modelización de variables discretas*. Universidad Autónoma de Madrid. Departamento de Economía Aplicada.

Mensah Y.M. (1983). The Differential Bankruptcy Predictive Ability of Specific Price Level Adjustments: Some Empirical Evidence. *The Accounting Review*, 53, 2, 228-246.

Mensah, Y.M. (1984). An examination of the stationarity of multivariate bankruptcy prediction models: A methodological study. *Journal of Accounting Research*, 22, 1, 380-395.

Messier, Jr.,W. y J. Hansen (1988). Inducing rules for expert system development: An example using default and bankruptcy data. *Management Science*, 34, 12, 1403-1415.

Meyer, P. y Pifer, H. (1970). Prediction of bank failures. *Journal of Finance*, 25, 4, 853-868.

Min, J.H. y Y.C. Lee (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28, 128–134.

Mínguez Conde, J.L. (2006). El fracaso empresarial en la empresa constructora. *Tesis Doctoral*. Universidad de Valladolid.

Moyer, R.Ch. (1977): Forecasting financial failure: A re-examination. *Financial Management*, 6, 1, 11-17.

Mulford, C.W. y E.E. Comiskey (2005). *Creative cash-flow reporting. Uncovering Sustainable Financial Performance*. John Wiley and Sons. New Jersey, 316 y ss.

Muller, G.H., Steyn-Bruwer y W.D. Hamman (2009). Predicting financial distress of companies listed on the JSE. A comparison of techniques. *South African Journal of Business Management*, 40, 1, 21-32.

Nasir, M.L., John, R.I., Bennett, S.C., Russell, D.M. y A. Patel (2000). Predicting corporate bankruptcy using artificial neural networks. *Journal of Applied Accounting Research*, 5, 3, 30-51.

Neophytou, E. y C.M. Molinero (2004). Predicting corporate failure in the UK: A multidimensional scaling approach. *Journal of Business Finance and Accounting*, 31, 5/6, 677-710.

Norton, C.L. y R.E. Smith (1979). A comparison of general price level and historical cost financial statements in the prediction of bankruptcy. *The Accounting Review*, 54, 1, 72-87.

Nour, M. (1994). Improved clustering and classification algorithms for the Kohonen self-organizing neural network. *PhD Dissertation*. Kent State University.

Odom, M.D. y R. Sharda (1990). A neural networks for bankruptcy prediction. *IEEE INNS International Joint Conference on Neural Networks*, 2, 17-21, 163-168

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18, 1, 109-131.

Olsen, M., Bellas, C. y L.V. Kish (1983). Improving the prediction of restaurant failure through ratio analysis. *International Journal of Hospitality Management*, 2, 187–193.

Olson, D. L., Delen, D. y Y. Meng (2012). Comparative analysis of data mining methods for bankruptcy prediction. *Decision Support Systems*, 52, 464- 473.

Pantalone, C. y M. Platt (1987a). Predicting commercial bank failure since deregulation. Federal Reserve Bank of Boston New England Economic Review (July/ August), 37-47.

Pantalone, C. y M. Platt (1987b). Predicting failure of savings and loan associations. *AREUEA Journal*, 15, 2, 46-64.

Park, S. M. Hancer (2012). A comparative study of logit and artificial neural networks in predicting bankruptcy in the hospitality industry. *Tourism Economics*, 18, 2, 311-338.

Park, C. S. y I. Han (2002). A case-based reasoning with the feature weights derived by analytic hierarchy process for bankruptcy prediction. *Expert Systems with Applications*, 23, 255–264.

Parsa H.G., Self J., Sydnr-Busso S. y H.J. Yoon (2011). Why Restaurants Fail? Part II. The Impact of Affiliation, Location, and Size on Restaurant Failures: results from a survival analysis. *Journal of Food Service Business Research*, 14, 360-370.

Patterson, D. (2001). Bankruptcy prediction: A model for the casino industry. *PhD Dissertation*. University of Nevada-Las Vegas.

Pendharkar, P.C. (2002). A Potential Use of Data Envelopment Analysis for the Inverse Classification Problem. *Omega*, 30, 3, 243-248.

Peña, D. (2002). *Análisis de datos multivariantes*. Ed. Mc Graw Hill. Madrid.

Pettway, R. y J. Sinkey, Jr.. (1980). Establishing on-site banking examination priorities: An early warning system using accounting and market information. *Journal of Finance*, 35,1, 137-150.

Platt, H.D. y M.B. Platt (1990). Development of a class of stable predictive variables: The case of bankruptcy prediction. *Journal of Business Finance and Accounting*, 17,1, 31-51.

Platt, H.D. y M.B. Platt (1991). A note in the use of industry-relative ratios in bankruptcy prediction. *Journal of Banking and Finance*, 15, 6, 1183-1194.

Platt, H.D. y M.B. Platt (2008). Financial distress comparison across three global regions. *Journal of Risk and Financial Management*, 1, 1, 129-162.

Pindado, J. Rodrigues, L. y Ch. De la Torre (2008). Estimating Financial Distress Likelihood. *Journal of Business Research*, 61, 995-1003.

Parsa, H. G., Self, J. T., Njite, D. y T. King (2005). Why restaurants fail?. *Cornell Hotel and Restaurant Administration Quarterly*, 46, 304-22.

Premachandra, I.M., Chen, Y. y J. Watson (2011). DEA as a tool for predicting corporate failure and success: a case of bankruptcy assessment. *Omega*, 39, 6, 620–626.

Rafiei, F.M., Manzari, S.M. y S. Bostanian (2011). Financial health prediction models using artificial neural networks, genetic algorithm and multivariate discriminant analysis: iranian evidence. *Expert Systems with Applications*, 38, 10210–10217.

Ravisankar, P., Ravi, V. y I. Bose (2010). Failure prediction of dotcom companies using neural network-genetic programming hybrids. *Informatic Science*, 180, 1257–1267.

Ravisankar, P. y V. Ravi (2010). Financial distress prediction in banks using Group Method of Data Handling neural network, counter propagation neural network and fuzzy ARTMAP. *Knowledge-Based Systems*, 23, 823–831.

Real Decreto Ley 1514/2007 de 16 de Noviembre. Plan General de Contabilidad.

Real Decreto Ley 1515/2007 de 16 de Noviembre. Plan General de Contabilidad de pequeñas y medianas empresas.

Real Decreto Ley 3/2009 de 27 de Marzo, de medidas urgentes en materia tributaria.

Rose, P. y J. Kolari (1985). Early warning systems as a monitoring device for bank condition. *Quarterly Journal of Business and Economics*, 24,1, 43-60.

Ryu, Y. y W. Yue (2005). Firm bankruptcy prediction: experimental comparison of isotonic separation and other

classification approaches. *IEEE Trans. Syst. Man Cybern–Part A: Systems Humans*, 21, 4, 265–276.

Salchenberger, L., Cinar, E. y N. Lash (1992). Neural networks: A new tool for predicting bank failures. *Decision Sciences*, 23, 899-916.

Sánchez, J.J. (1984). *Introducción a las técnicas de análisis multivariante aplicadas a las Ciencias Sociales*. Centro de Investigaciones Sociológicas, Madrid.

Sanchis, A. (2000). Una aplicación del análisis discriminante a la previsión de la insolvencia en las empresas españolas de seguros no-vida. *Tesis doctoral*. Universidad Complutense de Madrid. Facultad de Ciencias Económicas y Empresariales.

Santomero, A. y J. Vinso (1977). Estimating the probability of failure for comercial banks and the banking system. *Journal of Banking and Finance*, 1, 2, 185-205.

Sarkar, S. y R.S. Sriram (2001). Bayesian models for early warning of bank failures. *Management Science*, 47, 11, 1457–1475.

Schwarz, G. (1978). Estimating the dimension of a model, *Annals of Statistics*, 6, 461-464.

Scherrer, P.S. y T.J. Mathison (1996). Investment strategies for REIT inventories. *Real State Review*, 26, 1, 5-10.

Shah, J. y M. Murtaza (2000). A neural network model based clustering procedure for bankruptcy prediction. *American Business Review*, 18, 2, 80-86.

Sangjae Lee y W.S. Choi (2013). A multi-industry bankruptcy prediction model using back propagation neural network and multivariate discriminant analysis. *Expert Systems with Applications*, 40, 2941-2946.

Sean, S., Hyun, S. y W. Kim (2011). Dimensions of brand equity in the Chain Restaurant Industry. *Cornell Hospitality Quarterly*, 52, 429-437.

Serrano-Cinca, C. (1996): Self Organizing neural networks for financial diagnosis. *Decision Support Systems*, 17, 3, 227-238.

Serrano Cinca, C. (1997). Feedforward neural networks in the classification of financial information. *European Journal of Finance*, 1997, 3, 3, 183-202.

Serrano-Cinca, C. y B. Gutiérrez-Nieto (2013). Partial Least Square Discriminant Analysis for bankruptcy prediction. *Decision Support Systems*, 3, 1245-1255.

Shin, K-S, y Y.J. Lee (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications*, 23, 3, 321-328.

Shin, K. S., Lee, T. S. y H.J. Kim (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, 28, 127-135.

Shuk-Wern Ong, Voon Choong Yap y R.W.L. Khong (2011). Corporate failure prediction: a study of public listed companies in Malaysia. *Managerial Finance*, 37, 6, 553-564.

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business*, 74, 1, 101-124.

Sinkey, Jr., J. (1975). A multivariate statistical analysis of the characteristics of problematic banks. *Journal of Finance*, 30, 1, 21-36.

Skogsviki, K. (1980). Current cost accounting ratios as predictors of business failure. The Swedish case. *Journal of Business, Finance and Accounting*, 17, 1, 137-160.

Smith, M. y D.K. Liou (2007). Industrial sector and financial distress. *Managerial Auditing Journal*, 22, 4, 376-391.

Stone, M. (1974). Cross-validation choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B*, 36, 111-147.

Stroe, R. y N. Barbuta-Misu (2010). Predicting the financial performance of the building sector enterprises-Case Study of Galati County (Romania), *The Review of Finance and Banking*, 2, 1, 29-39.

Soo Cheong J., Park, K. y E. Lee-Ji (2012). Estimating Cash flow-at-Risk (CFaR). A Comparables Approach for

Restaurant Firms. *Cornell Hospitality Quarterly*, 52, 232-240.

Spicka, J. (2013). The financial condition of the construction companies before bankruptcy. *European Journal of Business and Management*, 5, 23, 160-169.

Susicky, J. (2011). Applicability of Bankruptcy Models at Agricultural Companies. *Acta VSFS*, 3/2011, 5, 241-252.

Sun, T., Chang, N. y G. Lee (1999). Dynamics of modeling in data mining. Interpretive approach to bankruptcy prediction. *Journal of Management Information Systems*, 16, 1, 63-85.

Sun, J. y H. Li (2008). Listed companies financial distress prediction based on weighted majority voting combination of multiple classifiers. *Expert Systems with Applications*, 35, 818–827.

Sun, L. y P. Shenoy (2007). Using Bayesian networks for bankruptcy prediction: Some methodological issues. *European Journal of Operations Research*, 180, 738–753.

Sun, J., Li, H., Huang, Q.H. y K.Y. He (2014). Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches. *Knowledge-Based Systems*, 57, 41–56.

Takahashi, K., Kurokawa, Y. y K. Watase (1984). Corporate bankruptcy prediction in Japan. *Journal of Banking and Finance*, 8, 2, 229-247.

Takane, Y. (1987). Introduction to special section. *Psychometrika*, 52, 3, 316-316.

Taffler, R. (1984). Empirical models for the monitoring of UK corporations. *Journal of Banking and Finance*, 8, 2, 199-227.

Tam, K. (1991). Neural network models and the prediction of bank bankruptcy. *Omega*, 19, 5, 429-445.

Tam, K. y M. Kiang (1992). Managerial applications of neural networks: the case of bank failure prediction. *Management Science*, 38, 7, 926-947.

Theodossiou, P. (1991). Alternative models for assessing the financial condition of Business in Greece. *Journal of Business and Accounting*, 18, 5, 697-720.

Theodossiou, P. (1993). Predicting Shifts in the Mean of a Multivariate Time Series Process: An Application in Predicting Business Failures. *Journal of the American Statistical Association*, 88, 422, 441-449.

Treewichayapong, S., Chunhachinda, P. y C. Padungsaksawasdi (2011). Bankruptcy Prediction of Real Estate Firms in Thailand, *The International Journal of Finance*, 23, 1, 6672-6691.

Tsai, C. (2009). Feature selection in bankruptcy prediction. *Knowledge-Based Systems*, 22, 2, 120-127.

Tsai, C., Hsu, Y. y D.C. Yen (2014). A comparative study of

classifier ensembles for bankruptcy prediction. *Applied Soft Computing*, 24, 977–984.

Tseng, F., y Y. Hu (2010). Comparing four bankruptcy prediction models: Logit, quadratic, interval logit, neural and fuzzy neural networks. *Expert Systems with Applications*, 37, 1846-1853.

Tsukuda, J. y S. Baba (1994). Predicting Japanese Corporate Bankruptcy in terms of financial data using Neural Network. *Computers and Industrial Engineering*, 27, 445-448.

Tzong-Huei Lin (2009). A cross model study of corporate financial distress prediction in Taiwan: Multiple discriminant analysis, logit, probit and neural networks models, *Neurocomputing*, 72, 3507-3516.

Varela, J. y A. Rial (2008). *Estadística práctica para la investigación en Ciencias de la Salud*. Netbiblo.

Varetto, F. (1998). Genetic algorithms applications in the analysis of insolvency risk. *Journal of Banking and Finance*, 22, 1421–1439.

Vavrina, J., Hampel, D. y J. Janová (2013). New approaches for the financial distress classification in agribusiness. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, vol. LXI, 4, 1177-1182.

Visauta, B. (2003). *Análisis Estadístico con SPSS para Windows*, Vol. II. E. McGraw-Hill.

Wang, B. (2004). Strategy changes and internet firm survival. *PhD. Dissertation*. University of Minnesota.

Wang, Y., Wang, S. y K.K. Lai (2005). A new fuzzy support vector machine to evaluate credit risk. *IEEE Transactions on Fuzzy Systems*, 13, 6, 820–831.

Wang, C.M. y Y.F. Huang (2009). Evolutionary-based feature selection approaches with new criteria for data mining: A case study of credit approval data. *Expert Systems with Applications*, 36, 5900–5908.

Ward, T. (1994). An empirical study of the incremental predictive ability of Beaver's naive Operating Cash Flow measure using four-state ordinal models of financial distress. *Journal of Business Finance and Accounting*, 21, 4, 547-561.

Wasilewski, M. y M. Madra (2008). An assessment of the agricultural enterprises' solvency with the usage of the Altman Model. *Acta Oeconomica et informatica*, 2, 50-55.

Wilcox, J. W. (1973). A prediction of business failure using accounting data. *Journal of Accounting Research*, 11, 163-179.

Wilson R.L. y R. Sharda (1994). Bankruptcy prediction using neural networks. *Decision Support Systems*, 11, 545-557.

Wu, C.H., Tzeng, G.H., Goo, Y.J. y W.C. Fang (2007). A real-valued genetic algorithm to optimize the parameters

of support vector machine for predicting bankruptcy. *Expert System with Applications*, 32, 397-408.

Wu, D., Liang, L. y Z. Yang (2008). Analysing the financial distress of Chinese public companies using probabilistic neural networks and multivariate discriminate analysis. *Socio-Economic Planning Science*, 42, 3, 206–220.

Wu, W. (2011). Improving classification accuracy and causal knowledge for better credit decisions. *International Journal of Neural Systems*, 21, 4, 297–309.

Xu, M. y C. Zhang (2009). Bankruptcy prediction: the case of Japanese listed companies, *Review of Accounting Studies*, 14, 534-558.

Yang, Z. R., Platt, M. B. y H.D. Platt (1999). Probabilistic neural networks in bankruptcy prediction. *Journal of Business Research*, 44, 67–74.

Yeh, C. C., Chi, D. J. y M.F. Hsu (2010). A hybrid approach of DEA, rough set and support vector machines for business failure prediction. *Expert Systems with Applications*, 37, 1535–1541.

Yi-Chung Hu y F.M. Tseng (2005). Applying back propagation neural networks to bankruptcy prediction, *International Journal of Electronic Business Management*, 3, 2, 97-103.

Youn, H, y Z. Gu (2010a). Predict US restaurant firm failures: the artificial neural network mode versus logistic

regression model. *Tourism and Hospitality Research*, 10, 3, 171-187.

Youn, H, y Z. Gu (2010b). Predicting Korean lodging firm failures: An artificial neural network model along with a logistic regression model. *International Journal of Hospitality Management*, 29, 120-127.

Zanganeh, T., Rabiee, M. y M. Zarei (2011). Applying Neuro-Fuzzy Model for Bankruptcy Prediction. *International Journal of Computer Applications*, 20, 3, 15-21.

Zacharakis, A., Meyer, D. y J. De Castro (1999). Differing perceptions of new venture failure: A matched exploratory study of venture capitalists and entrepreneurs. *Journal of Small Business Management*, 37, 1-14.

Zhang, G., Hu, M.Y., Patuwo, B.E. y D.C. Indro (1999). Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operations Research*, 116, 1, 16-33.

Zhang, L., Zhang, L., Teng, W. y Ch. Yibing (2013). Based on Information Fusion Technique with Data Mining in the Application of Finance Early-Warning. *Procedia Computer Science*, 17, 695-703.

Zhang, Y. y Y. Yang (2015). Cross-validation for selecting a model selection procedure. *Journal of Econometrics*, 187, 95-112.

Zhou, L. (2013). Performance of corporate bankruptcy prediction models on imbalanced dataset: The effect of sampling methods. *Knowledge-Based Systems*, 41, 16–25.

Zavgren, C.V. (1985). Assessing the vulnerability to failure of American industrial firms. *Journal of Business and Accounting*, 12, 1, 19-45.

Zimmer, I. (1980). A lens study of the prediction of corporate failure by bank loan officers. *Journal of Accounting Research*, 18, 2, 629-636.

Zmijewski, M.E. (1984). Methodological Issues Related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22, 59-81.

ANEXOS

ANEXO 1: INDICE DE TABLAS

Tabla	Título	Pág.
1.1	Principales modelos descentrados para la predicción de la insolvencia	143
1.2	Principales modelos centrados (AGRICULTURA) para la predicción de la insolvencia	147
1.3	Principales modelos centrados (INDUSTRIA) para la predicción de la insolvencia	149
1.4	Principales modelos centrados (CONSTRUCCIÓN) para la predicción de la insolvencia	152
1.5	Principales modelos centrados (COMERCIO Y SERVICIOS) para la predicción de la insolvencia	154
1.6	Principales modelos centrados (HOSTELERÍA) para la predicción de la insolvencia	156
3.1	Frecuencia de utilización de los ratios más comunes en la literatura previa	249
3.2	Descripción de las variables utilizadas	251
3.3	Definición de las partidas incluidas en las variables utilizadas	253
4.1	Estadísticos descriptivos MODELO GLOBAL M.I	267
4.2	Análisis de normalidad MODELO GLOBAL M.I	270
4.3	Tests no paramétricos MODELO GLOBAL M.I	271
4.4	Análisis de correlación bilateral MODELO GLOBAL M.I	273
4.5	MODELO LOGIT GLOBAL sin variables cualitativas M.I	280
4.6	Resultados posibles en una matriz de clasificación	281
4.7	Estadísticos descriptivos MODELO AGRICULTURA M.I	284
4.8	Análisis de normalidad MODELO AGRICULTURA M.I	285
4.9	Tests no paramétricos MODELO AGRICULTURA M.I	286
4.10	Análisis de correlación bilateral MODELO AGRICULTURA M.I	288
4.11	MODELO LOGIT AGRICULTURA M.I	291
4.12	Estadísticos descriptivos MODELO INDUSTRIA M.I	294
4.13	Análisis de normalidad MODELO INDUSTRIA M.I	295
4.14	Tests no paramétricos MODELO INDUSTRIA M.I	297
4.15	Análisis de correlación bilateral MODELO INDUSTRIA M.I	299

Tabla	Título	Pág.
4.16	MODELO LOGIT INDUSTRIA M.I	301
4.17	Estadísticos descriptivos MODELO CONSTRUCCIÓN M.I	304
4.18	Análisis de normalidad MODELO CONSTRUCCIÓN M.I	305
4.19	Tests no paramétricos MODELO CONSTRUCCIÓN M.I	306
4.20	Análisis de correlación bilateral MODELO CONSTRUCCIÓN M.I	308
4.21	MODELO LOGIT CONSTRUCCIÓN M.I	310
4.22	Estadísticos descriptivos MODELO COMERCIO Y SERVICIOS M.I	313
4.23	Análisis de normalidad MODELO COMERCIO Y SERVICIOS M.I	314
4.24	Tests no paramétricos MODELO COMERCIO Y SERVICIOS M.I	315
4.25	Análisis de correlación bilateral MODELO COMERCIO Y SERVICIOS M.I	317
4.26	MODELO LOGIT COMERCIO Y SERVICIOS M.I	319
4.27	Estadísticos descriptivos MODELO HOSTELERÍA M.I	322
4.28	Análisis de normalidad MODELO HOSTELERÍA M.I	323
4.29	Tests no paramétricos MODELO HOSTELERÍA M.I	324
4.30	Análisis de correlación bilateral MODELO HOSTELERÍA M.I	325
4.31	MODELO LOGIT HOSTELERÍA M.I	328
4.32	Modelos Global y Sectoriales obtenidos para M.I	330
4.33	MODELO LOGIT GLOBAL con variables cualitativas M.I	333
4.34	Comparación entre los modelos globales (Con y Sin inclusión de dummies) en M.I	335
4.35	Resultados de los Modelos de Regresión Logística para M.I	338
4.36	Robustez de la regresión logística del modelo global con variables cualitativas para M.I	339
4.37	Estadísticos descriptivos MODELO GLOBAL M.II	341
4.38	Análisis de normalidad MODELO GLOBAL M.II	342
4.39	Tests no paramétricos MODELO GLOBAL M.II	343

Tabla	Título	Pág.
4.40	Análisis de correlación bilateral MODELO GLOBAL M.II	344
4.41	MODELO LOGIT GLOBAL sin variables cualitativas M.II	346
4.42	Estadísticos descriptivos MODELO AGRICULTURA M.II	348
4.43	Análisis de normalidad MODELO AGRICULTURA M.II	349
4.44	Tests no paramétricos MODELO AGRICULTURA M.II	350
4.45	Análisis de correlación bilateral MODELO AGRICULTURA M.II	351
4.46	MODELO LOGIT AGRICULTURA M.II	353
4.47	Estadísticos descriptivos MODELO INDUSTRIA M.II	354
4.48	Análisis de normalidad MODELO INDUSTRIA M.II	355
4.49	Tests no paramétricos MODELO INDUSTRIA M.II	356
4.50	Análisis de correlación bilateral MODELO INDUSTRIA M.II	358
4.51	MODELO LOGIT INDUSTRIA M.II	360
4.52	Estadísticos descriptivos MODELO CONSTRUCCIÓN M.II	361
4.53	Análisis de normalidad MODELO CONSTRUCCIÓN M.II	362
4.54	Tests no paramétricos MODELO CONSTRUCCIÓN M.II	363
4.55	Análisis de correlación bilateral MODELO CONSTRUCCIÓN M.II	365
4.56	MODELO LOGIT CONSTRUCCIÓN M.II	367
4.57	Estadísticos descriptivos MODELO COMERCIO Y SERVICIOS M.II	369
4.58	Análisis de normalidad MODELO COMERCIO Y SERVICIOS M.II	370
4.59	Tests no paramétricos MODELO COMERCIO Y SERVICIOS M.II	371
4.60	Análisis de correlación bilateral MODELO COMERCIO Y SERVICIOS M.II	372
4.61	MODELO LOGIT COMERCIO Y SERVICIOS M.II	374
4.62	Estadísticos descriptivos MODELO HOSTELERÍA M.II	376
4.63	Análisis de normalidad MODELO HOSTELERÍA M.II	377
4.64	Tests no paramétricos MODELO HOSTELERÍA M.II	378

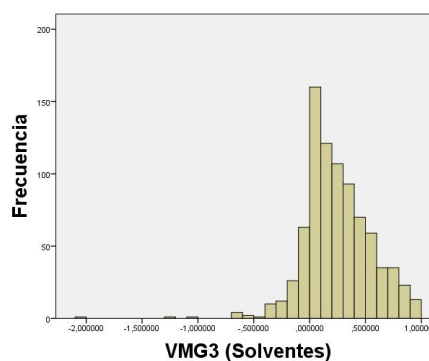
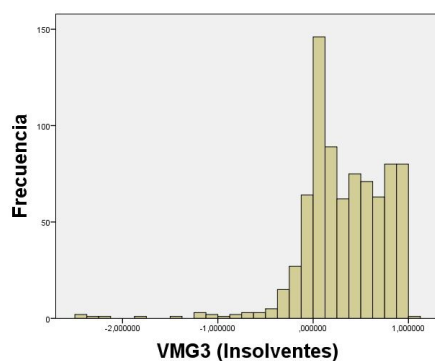
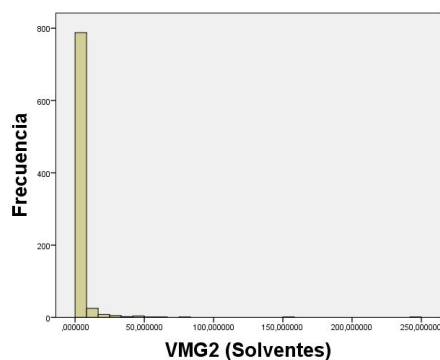
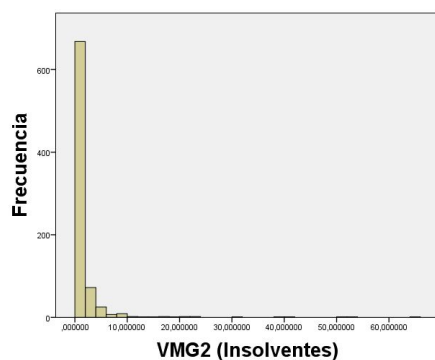
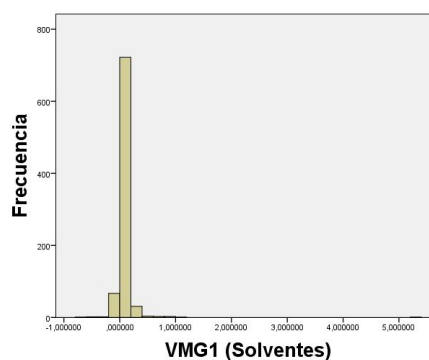
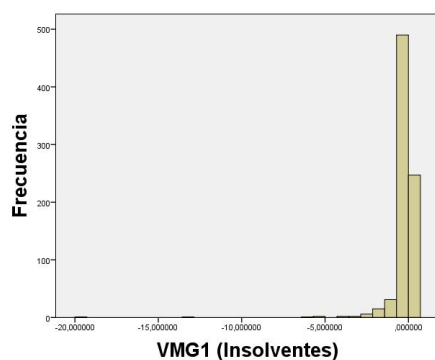
Tabla	Título	Pág.
4.65	Análisis de correlación bilateral MODELO HOSTELERÍA M.II	379
4.66	MODELO LOGIT HOSTELERÍA M.II	381
4.67	Modelos Global y Sectoriales obtenidos para M.II	383
4.68	MODELO LOGIT GLOBAL con variables cualitativas M.II	386
4.69	Comparación entre los modelos globales (Con y Sin inclusión de dummies) en M.II	388
4.70	Resultados de los Modelos de Regresión Logística para M.II	390
4.71	Robustez de la regresión logística del modelo global con variables cualitativas para M.I	391
4.72	Comparación entre los modelos globales y sectoriales en M.I y M.II	393
4.73	Comparación entre los modelos globales en M.I y M.II	394
4.74	Resultados de la comparación entre modelos globales y sectoriales	395

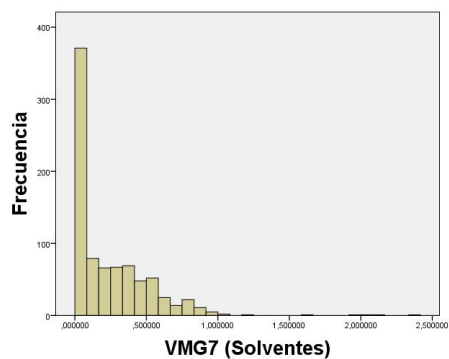
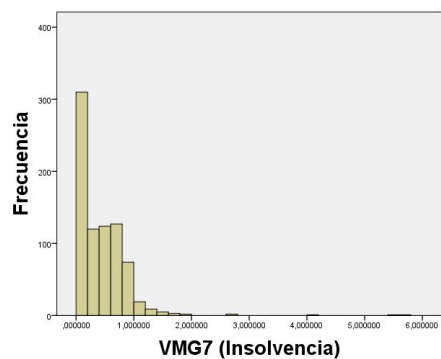
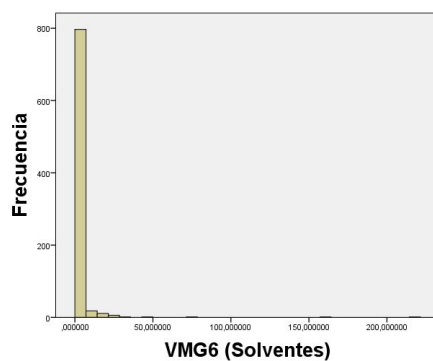
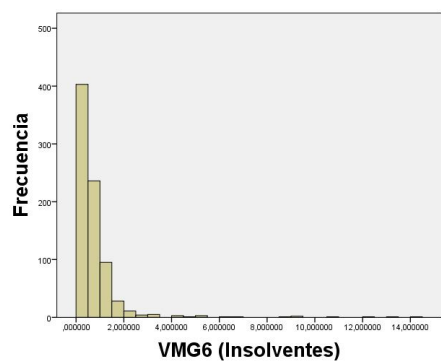
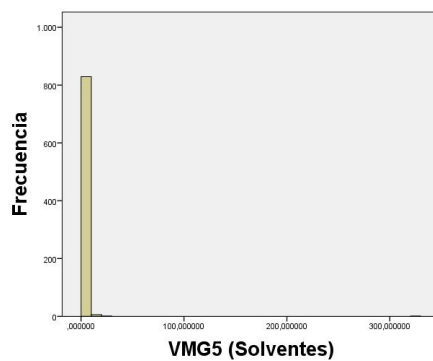
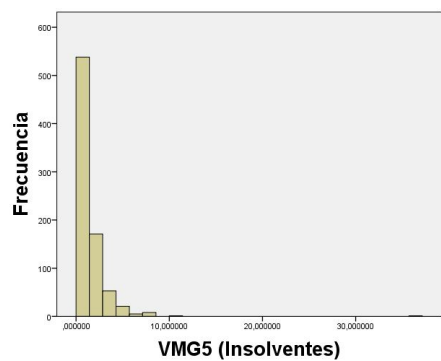
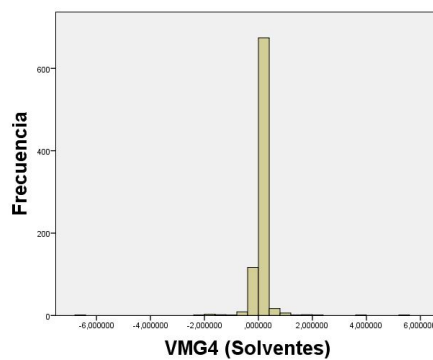
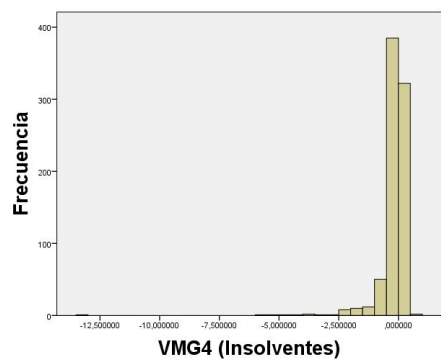
ANEXO 2: INDICE DE GRÁFICOS

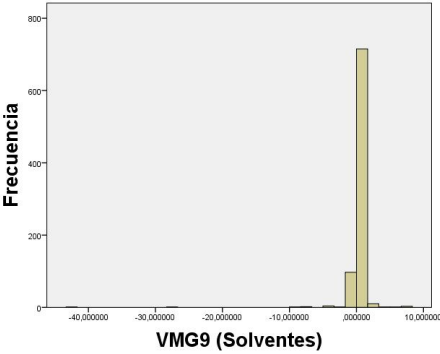
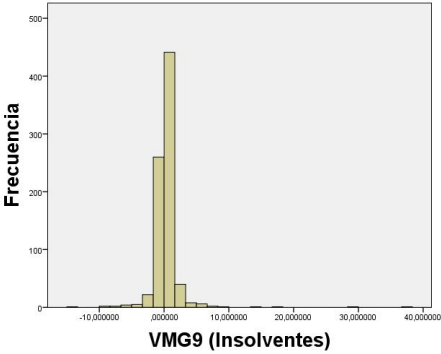
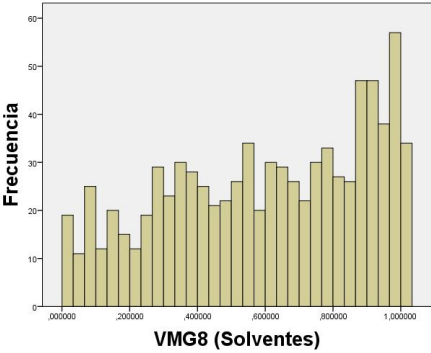
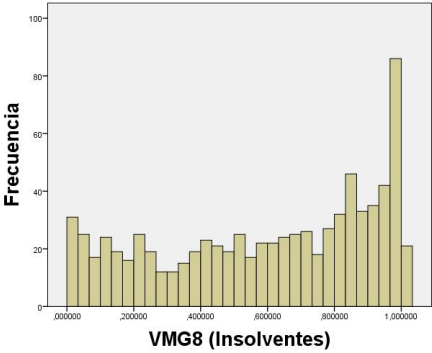
Gráfico	Título	Pág.
2.1	Ejemplos de funciones discriminantes	206
2.2	Clasificación de los modelos de respuesta cualitativa	210
2.3	Representación gráfica de los modelos LOGIT y PROBIT	214
3.1	Ficha técnica del muestreo	247
4.1	Metodología aplicada	264

ANEXO 3: HISTOGRAMAS

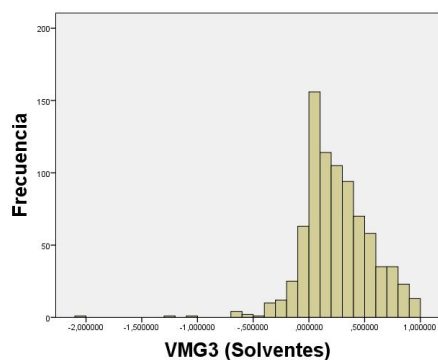
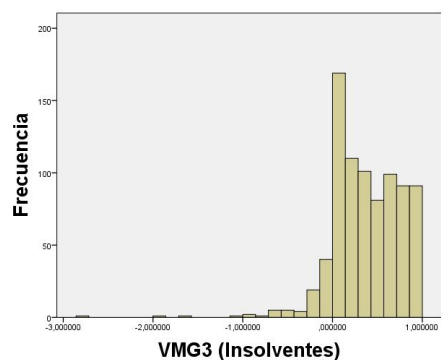
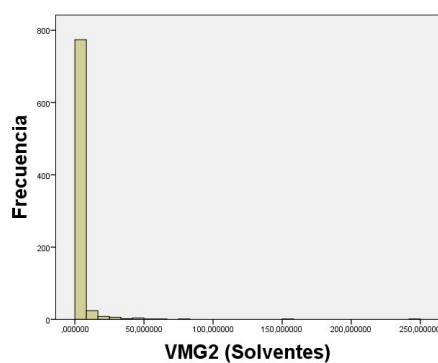
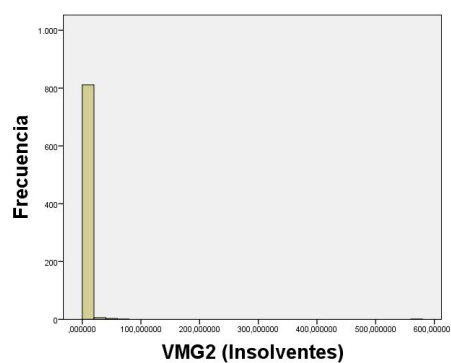
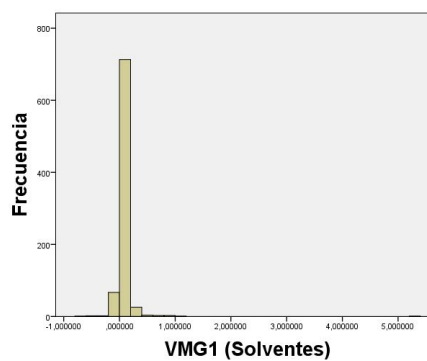
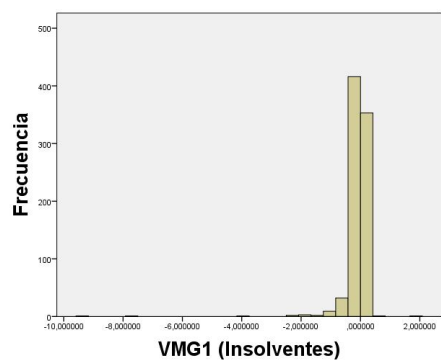
Histogramas de las variables de la Muestra Global, M.I

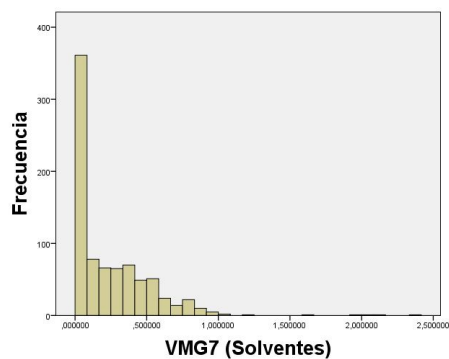
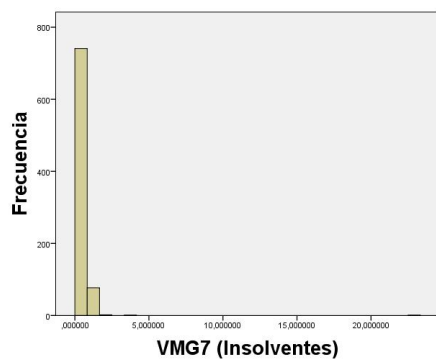
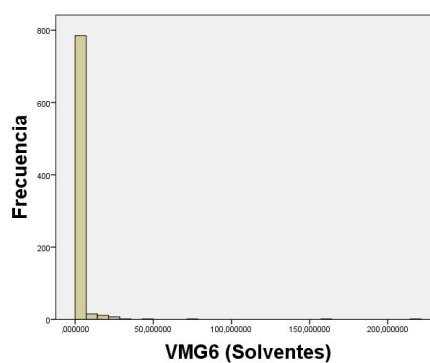
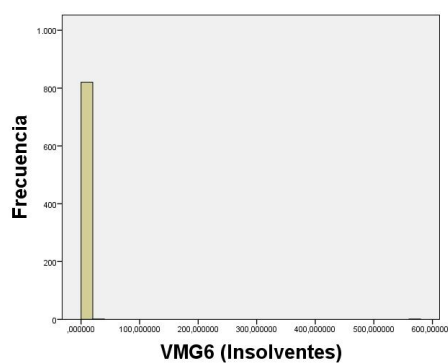
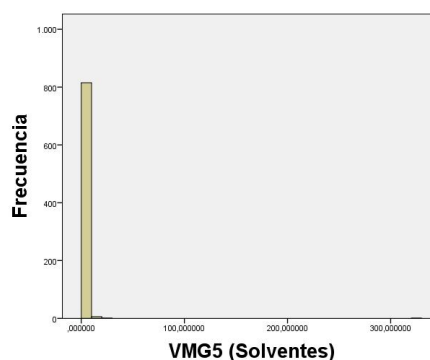
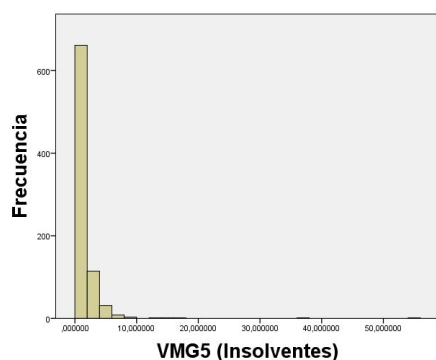
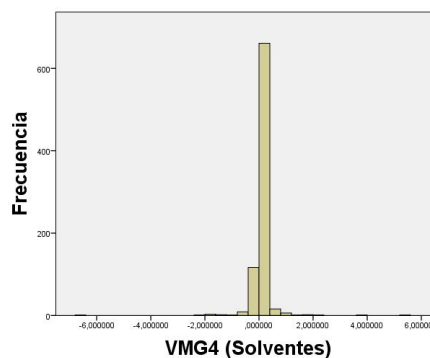
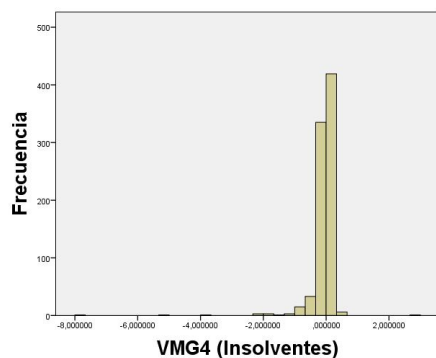


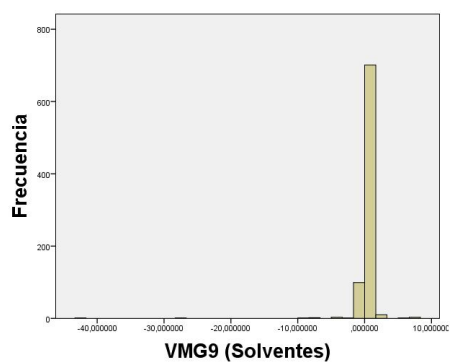
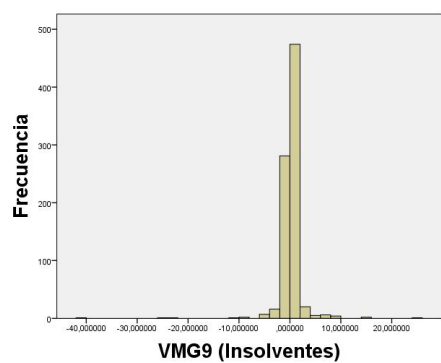
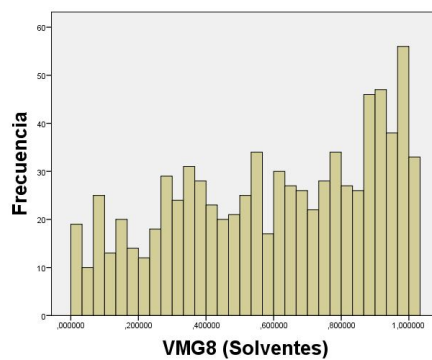
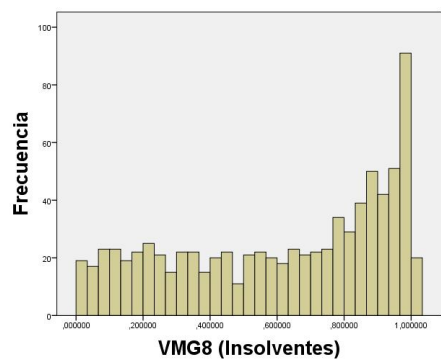




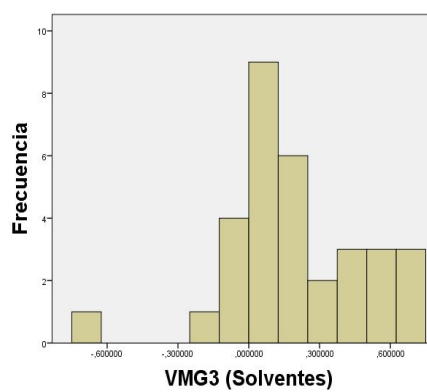
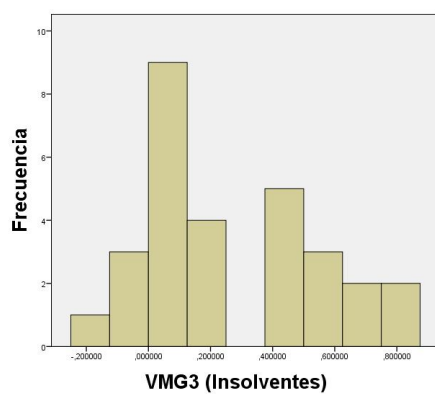
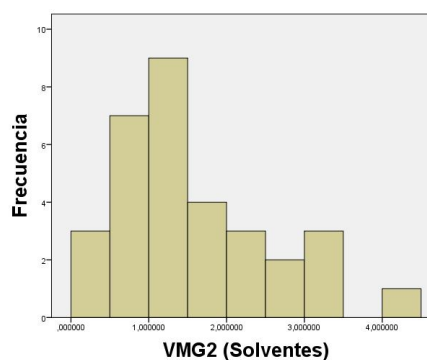
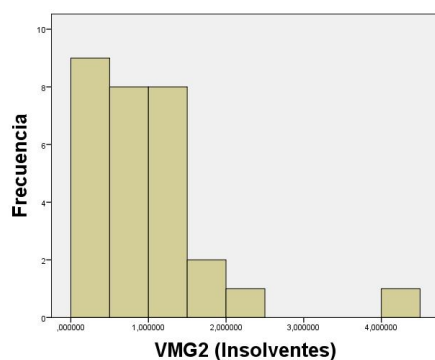
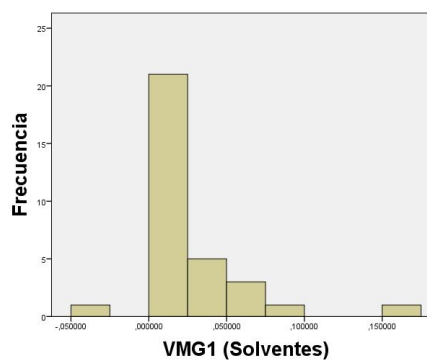
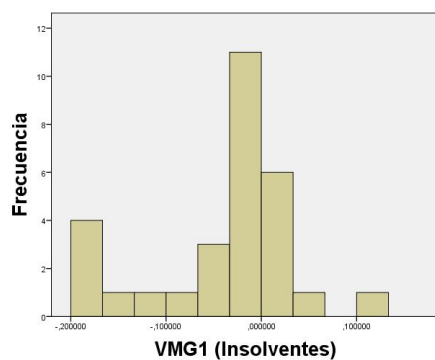
Histogramas de las variables de la Muestra Global, M.II

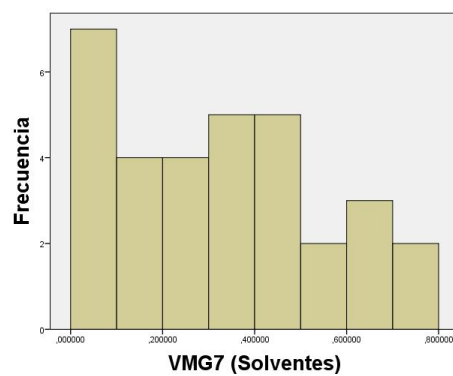
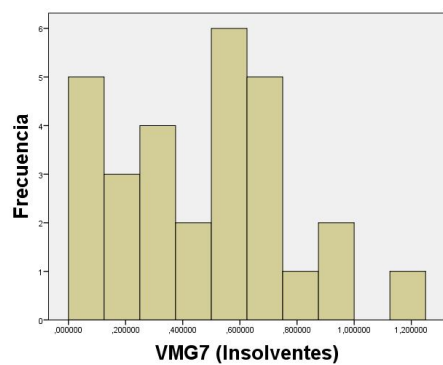
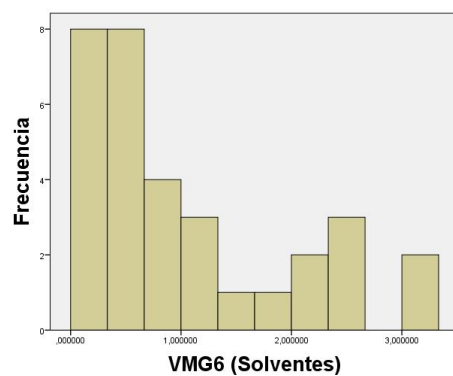
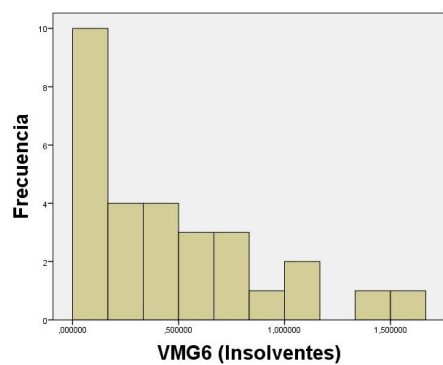
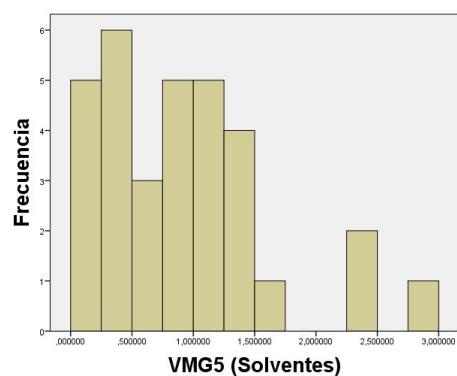
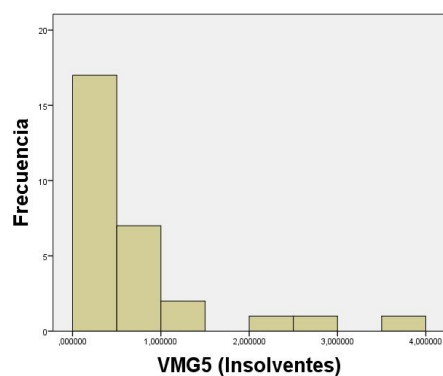
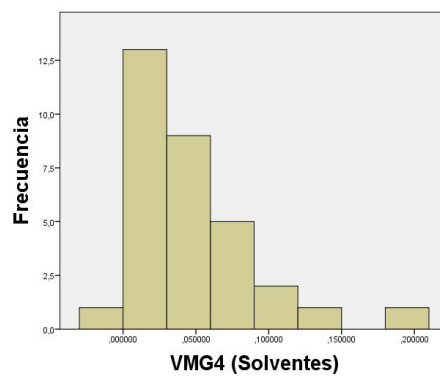
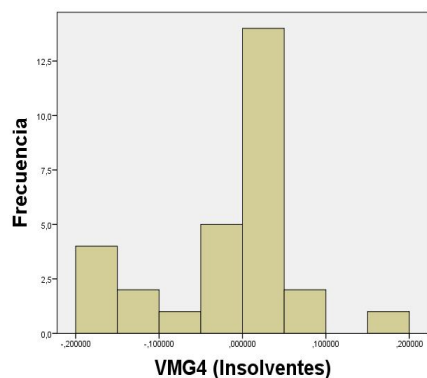


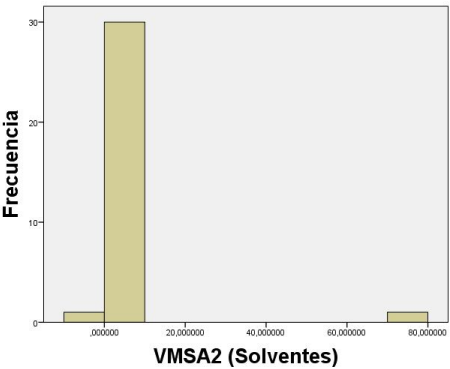
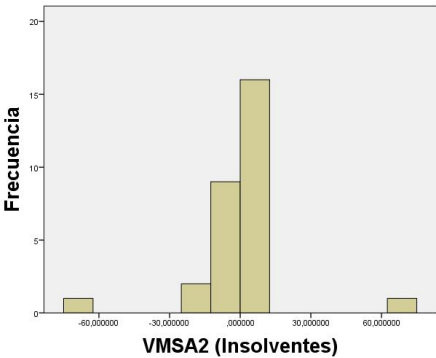
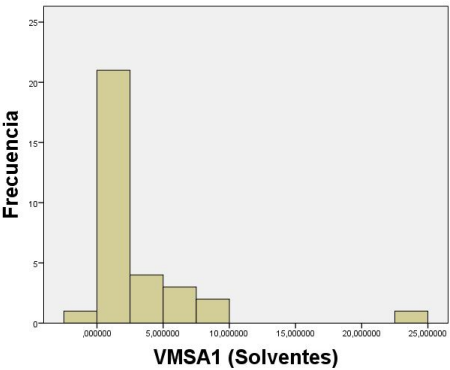
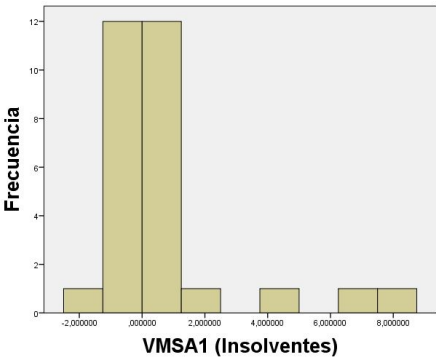
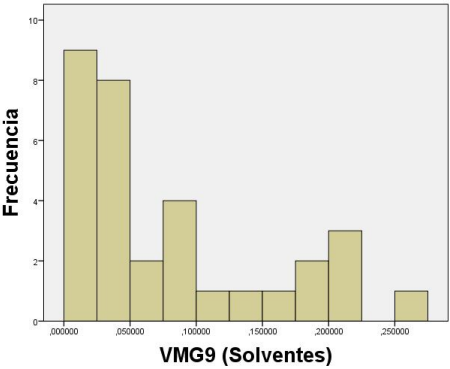
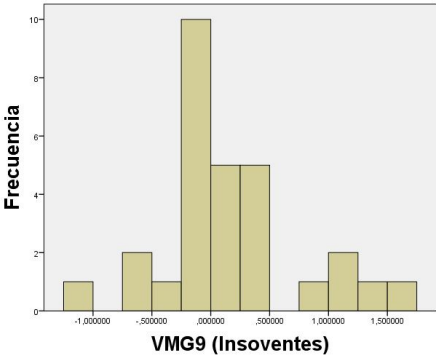
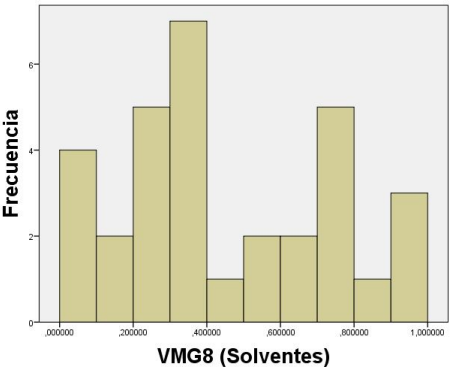
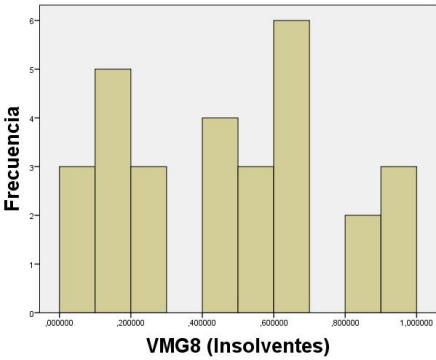


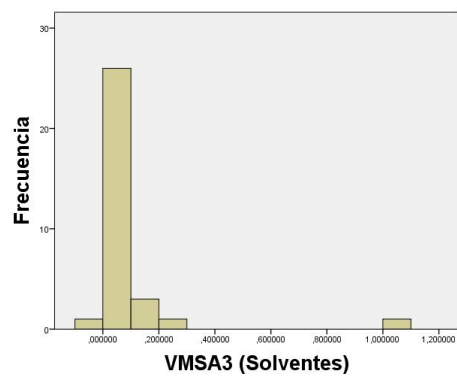
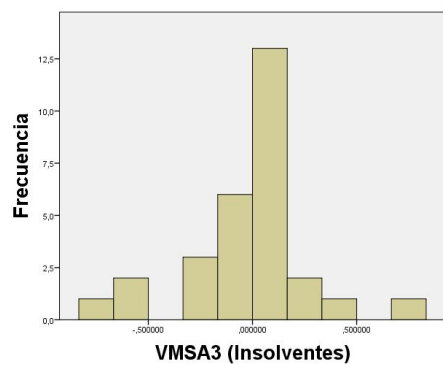


Histogramas de las variables del Sector Agricultura, M.I

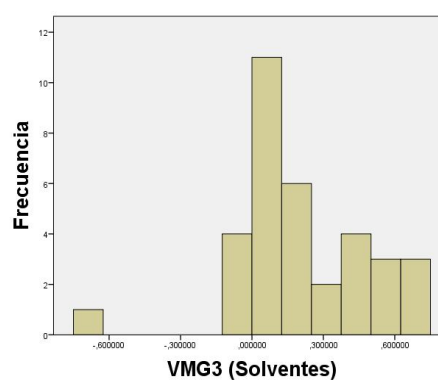
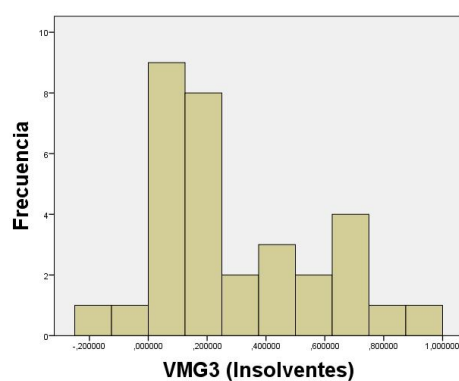
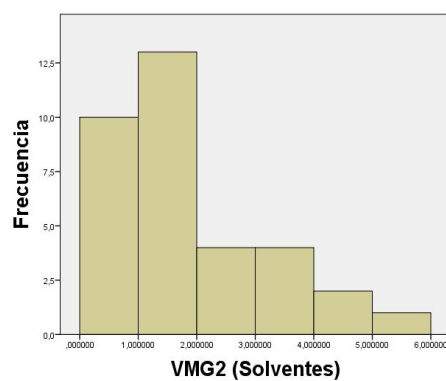
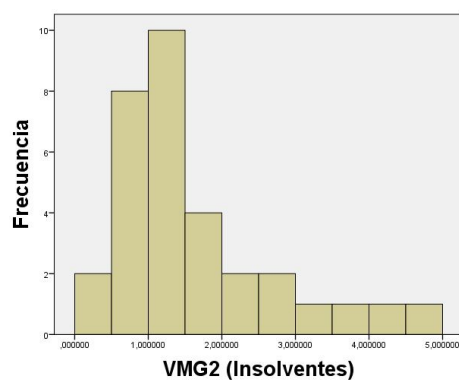
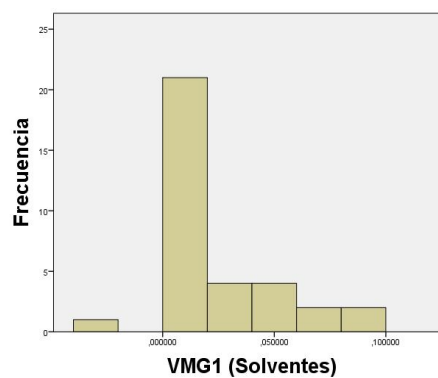
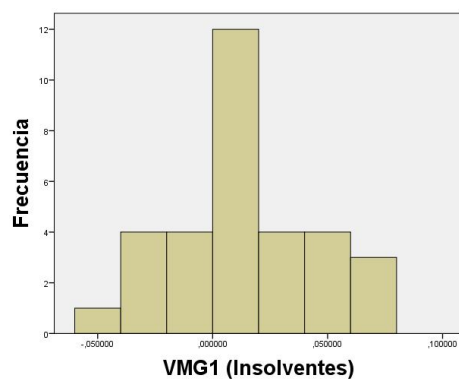


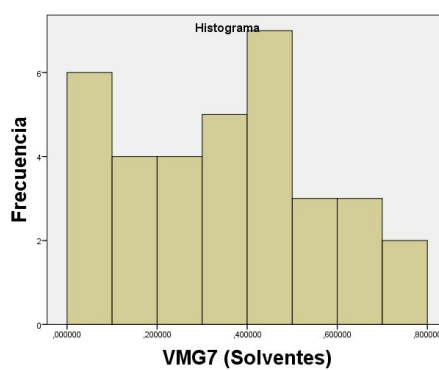
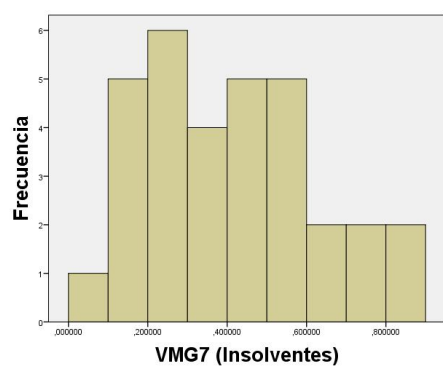
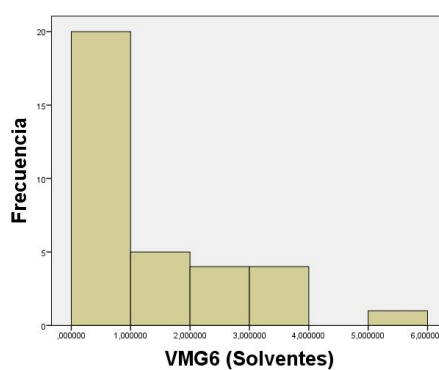
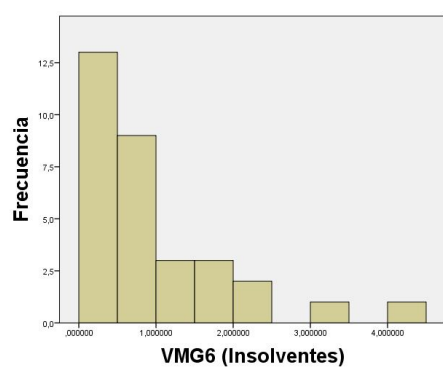
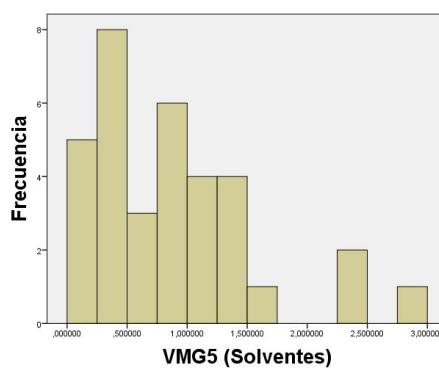
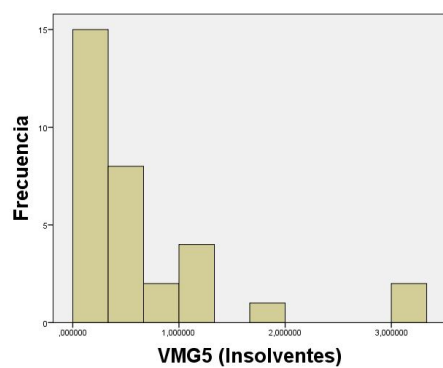
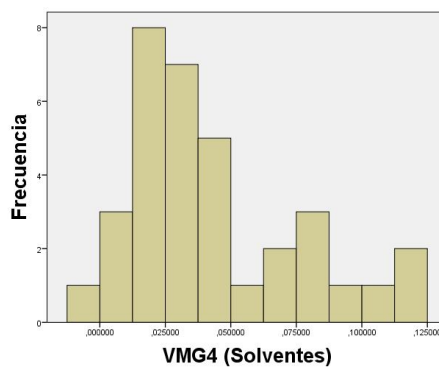
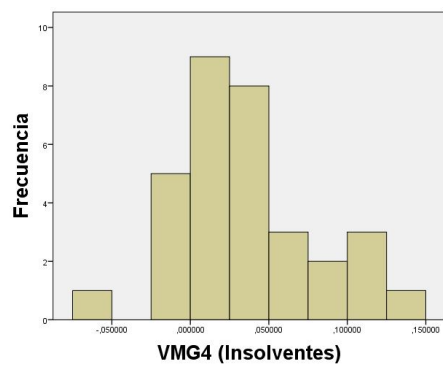


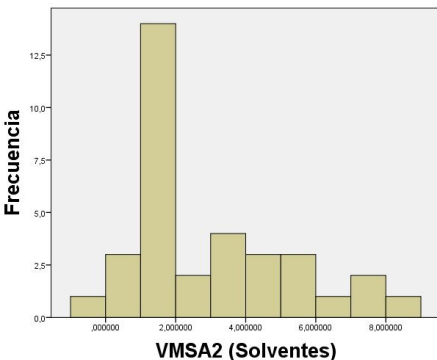
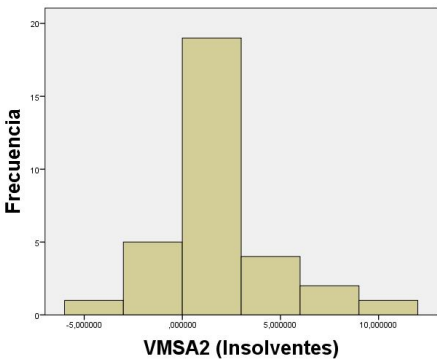
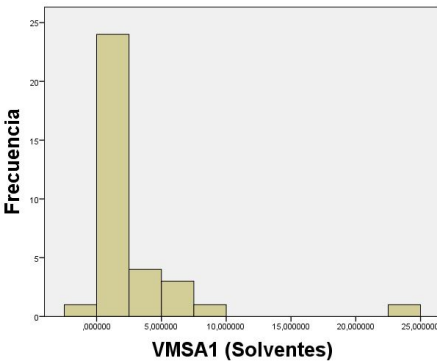
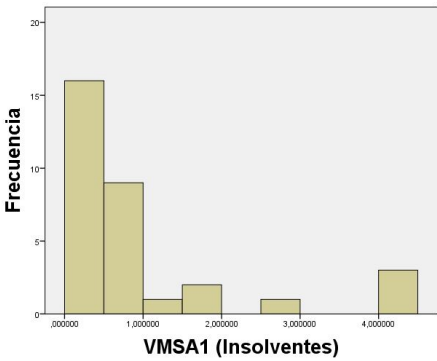
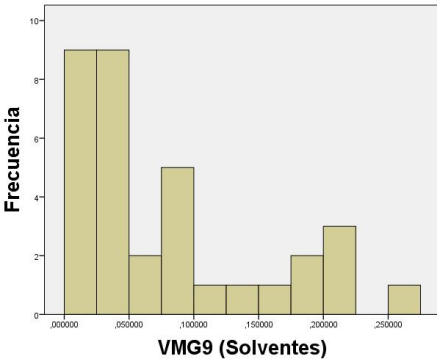
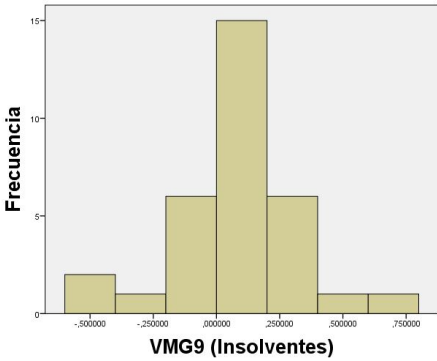
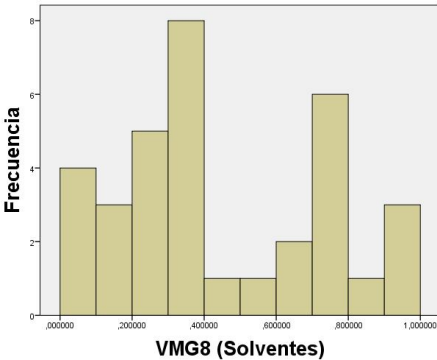
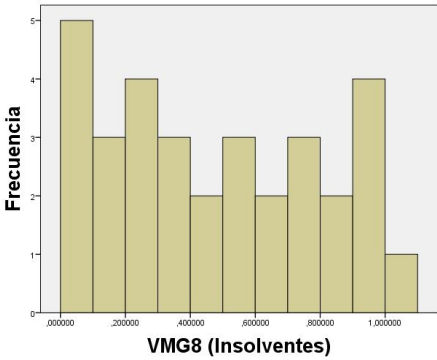


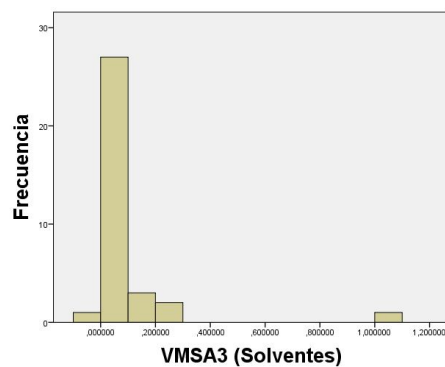
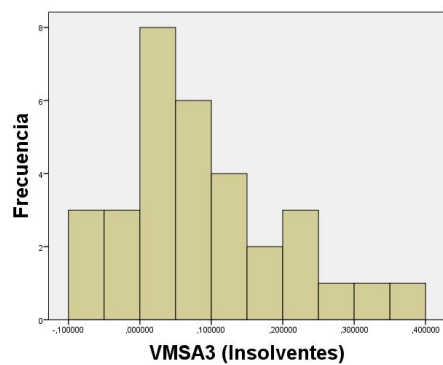


Histogramas de las variables del Sector Agricultura, M.II

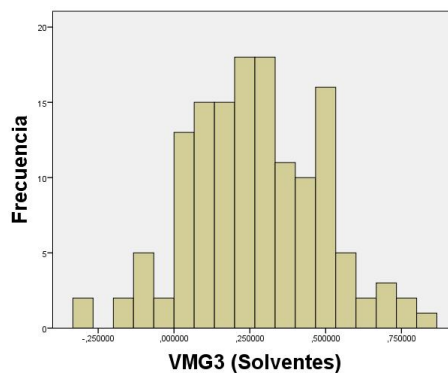
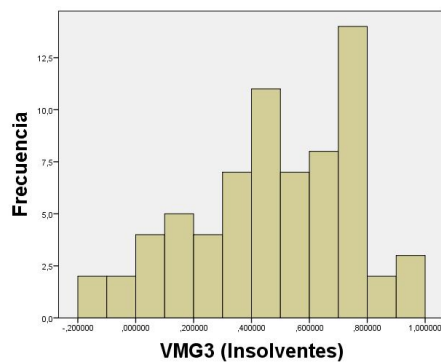
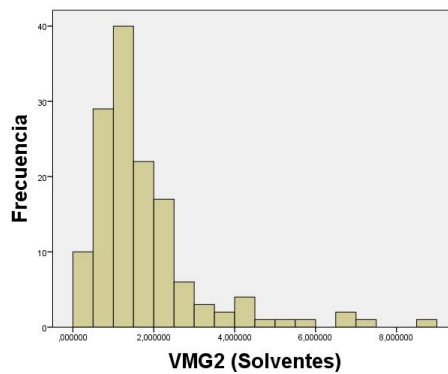
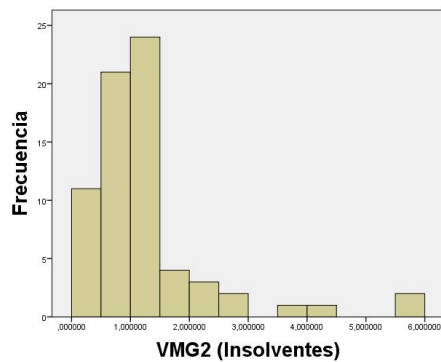
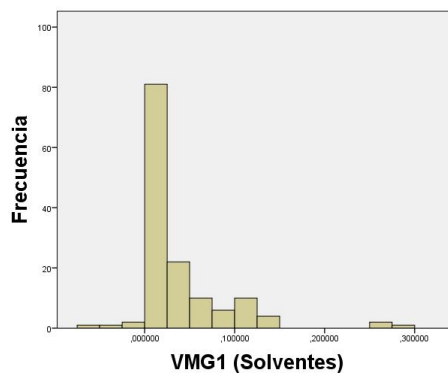
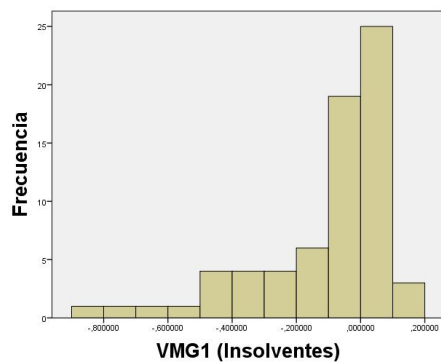


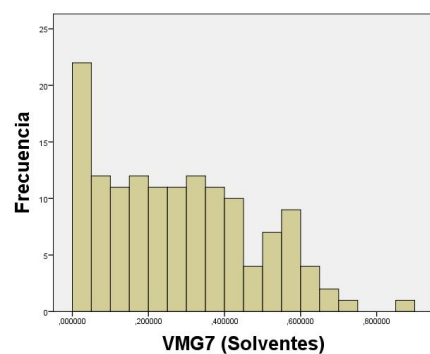
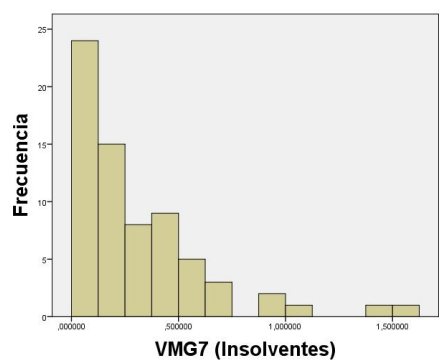
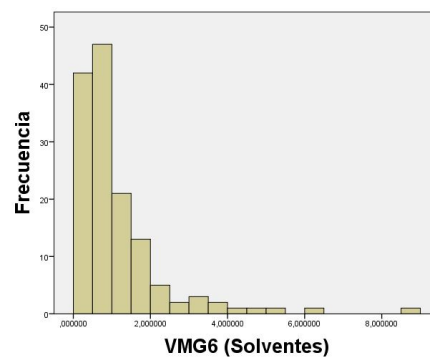
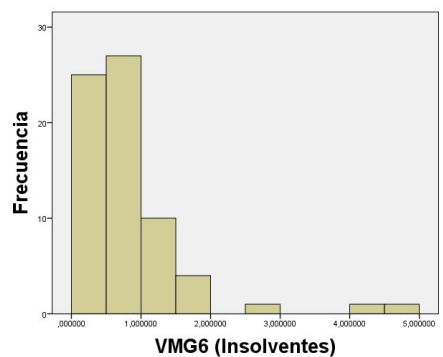
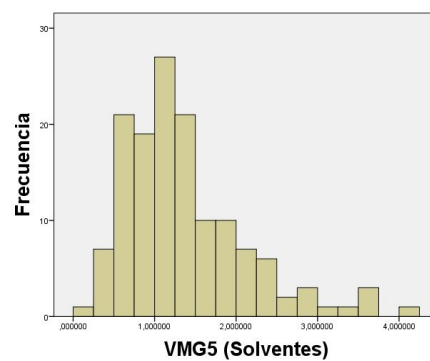
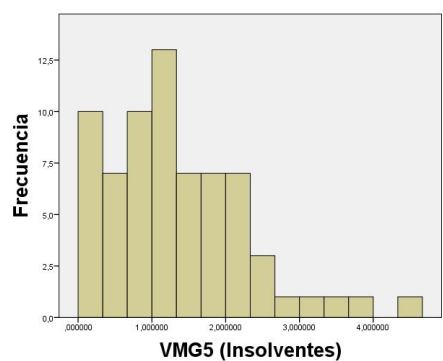
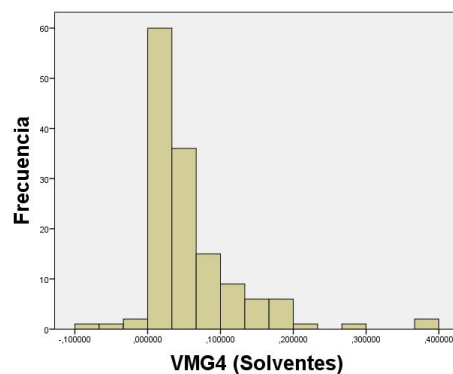
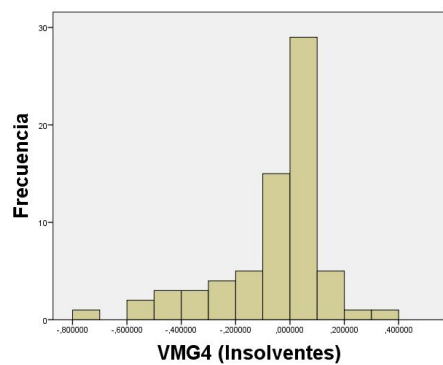


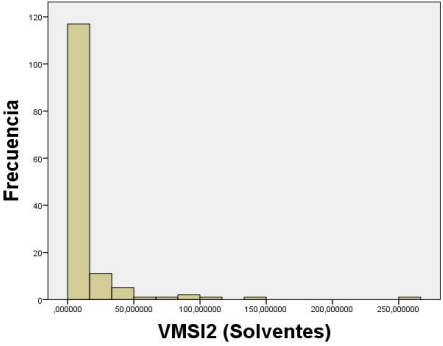
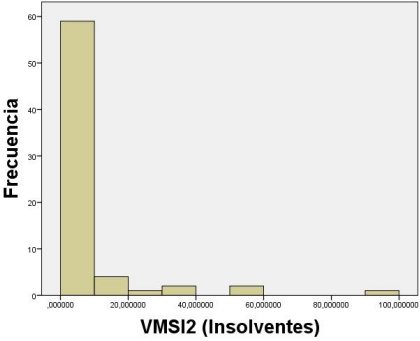
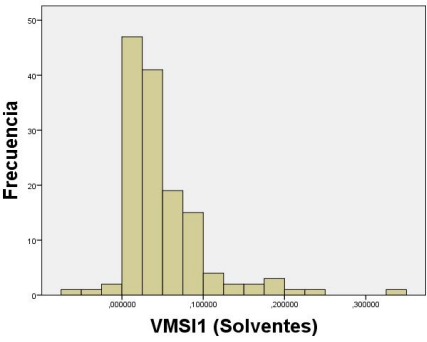
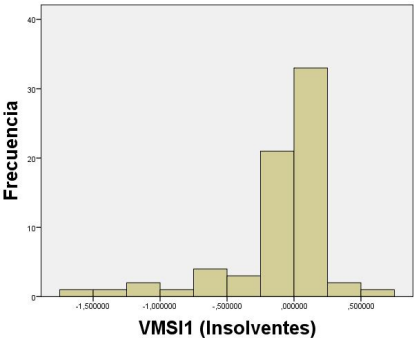
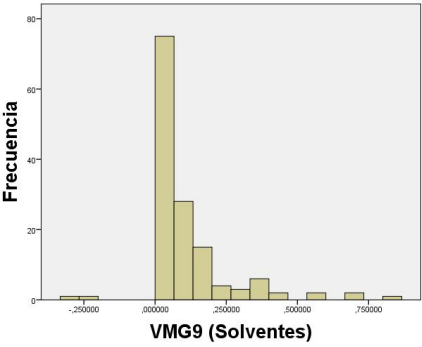
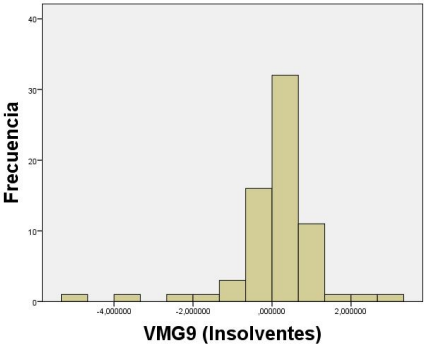
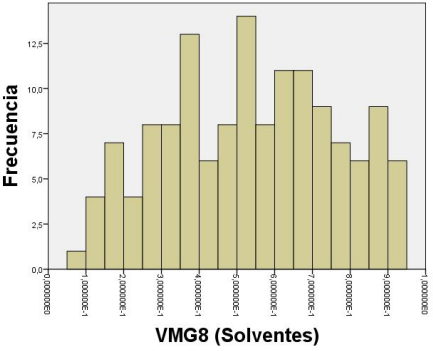
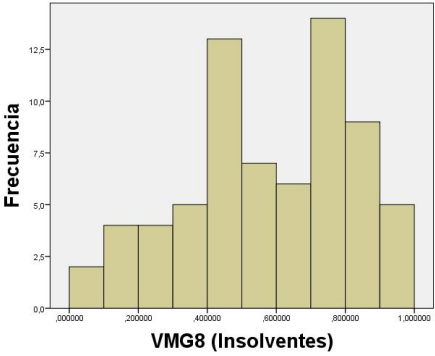


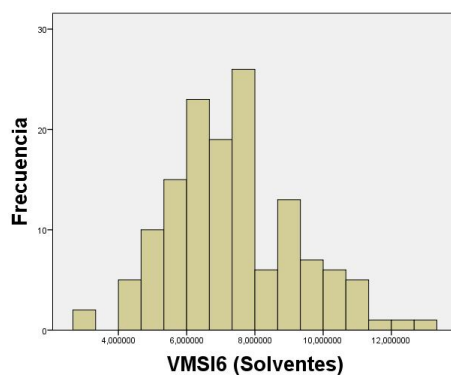
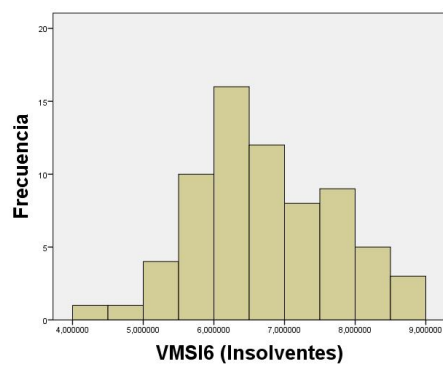
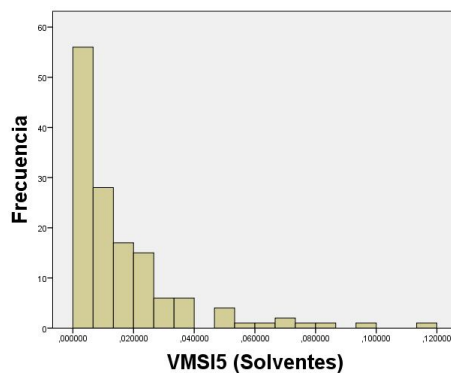
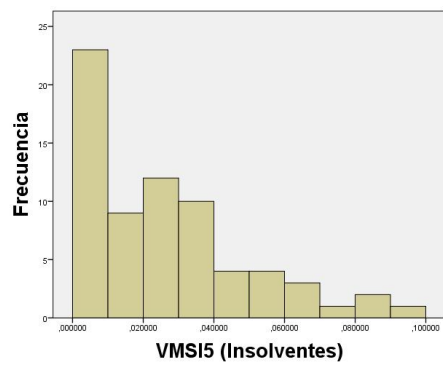
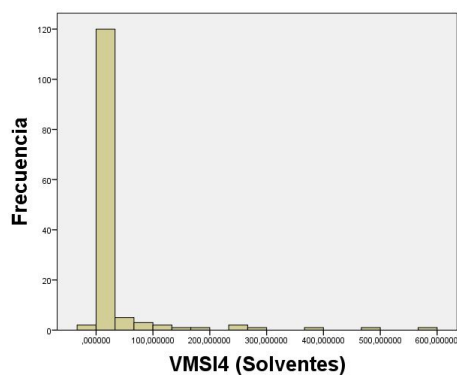
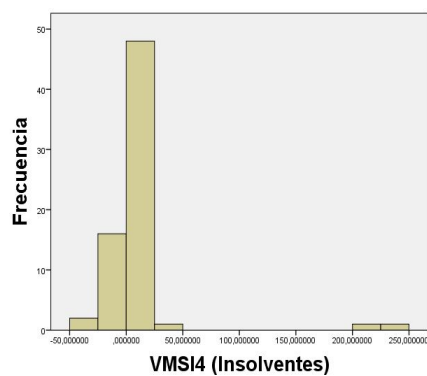
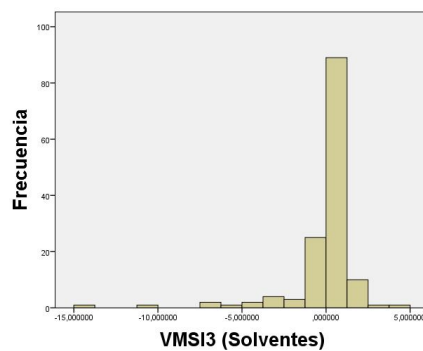
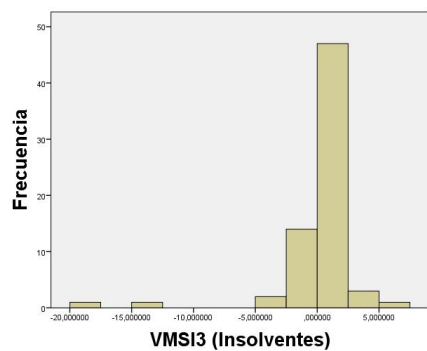


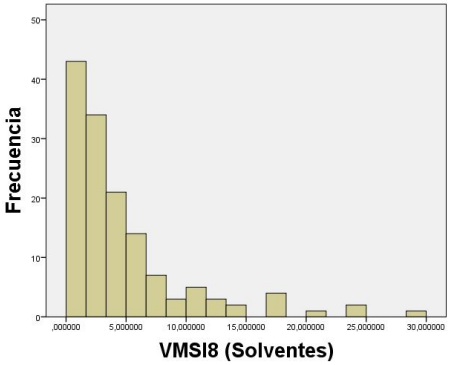
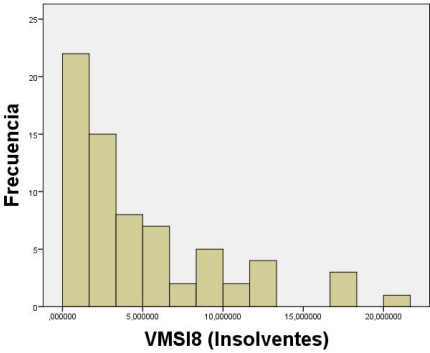
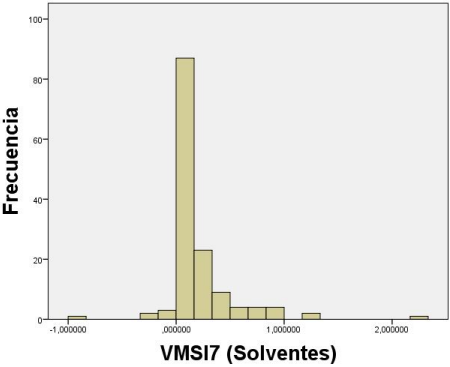
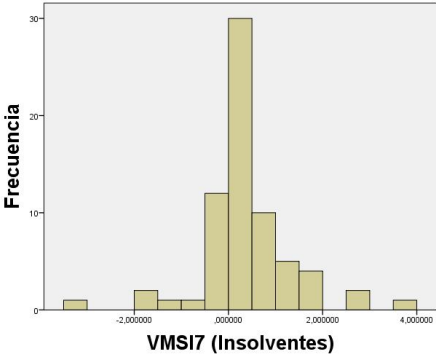
Histogramas de las variables del Sector Industria, M.I



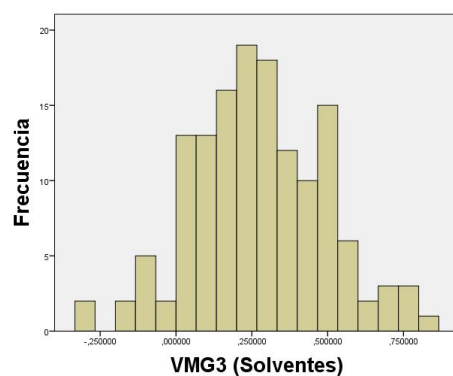
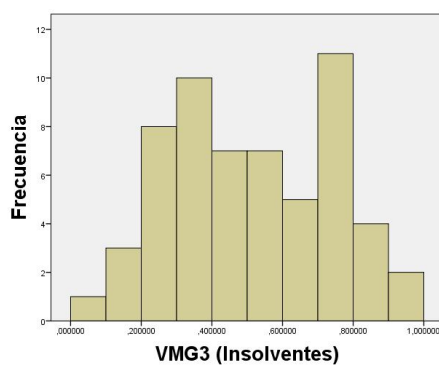
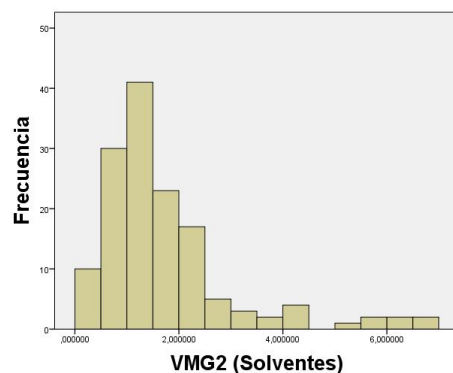
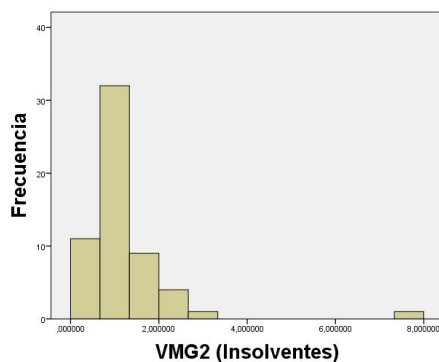
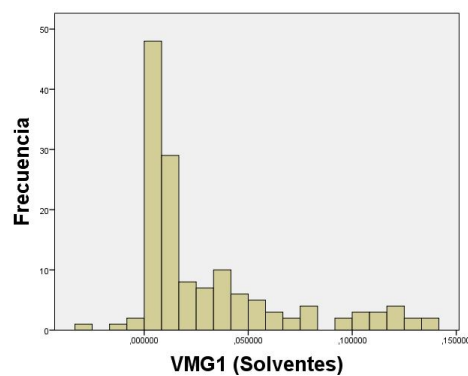
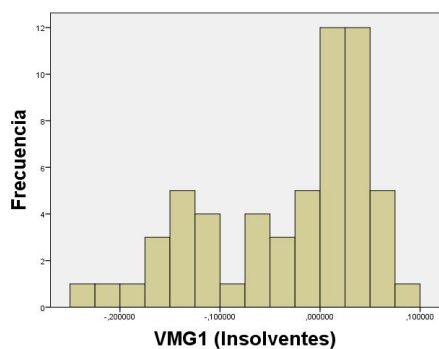


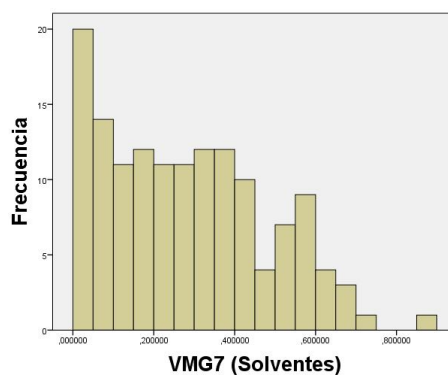
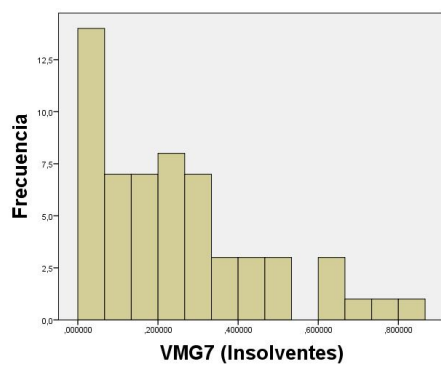
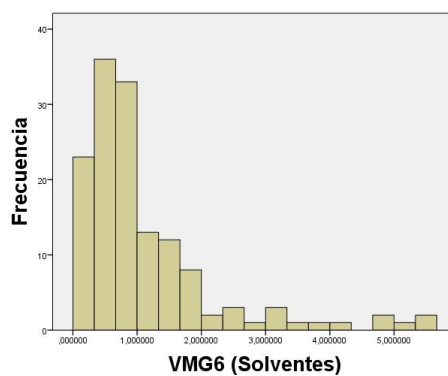
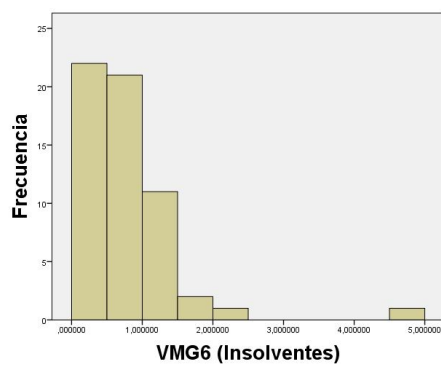
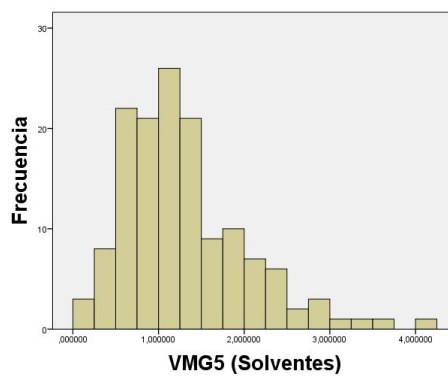
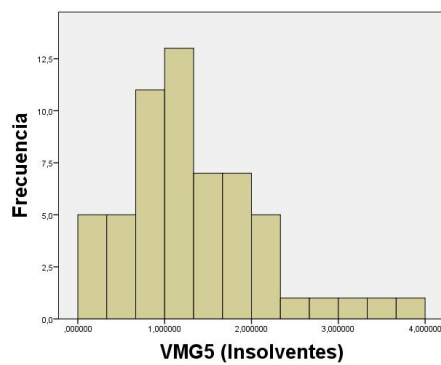
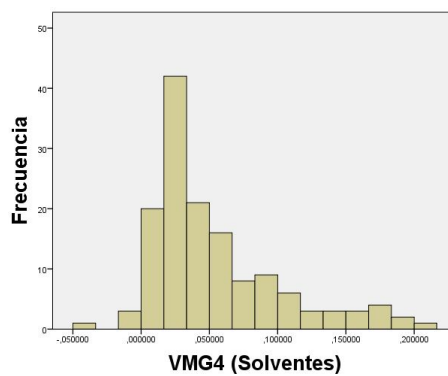
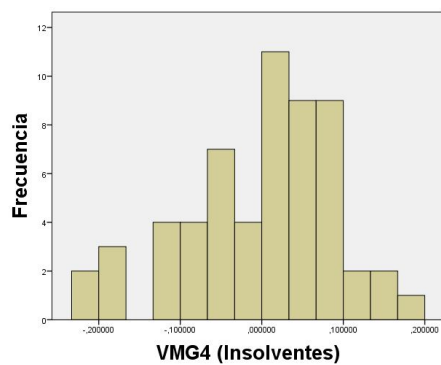


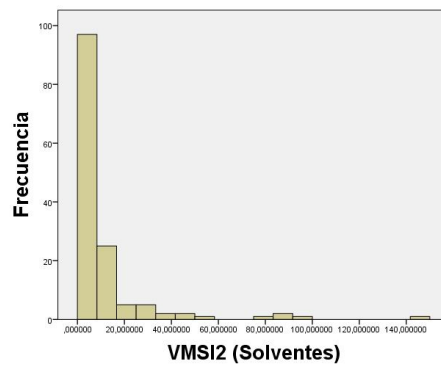
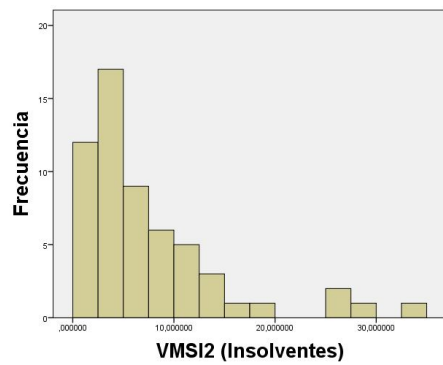
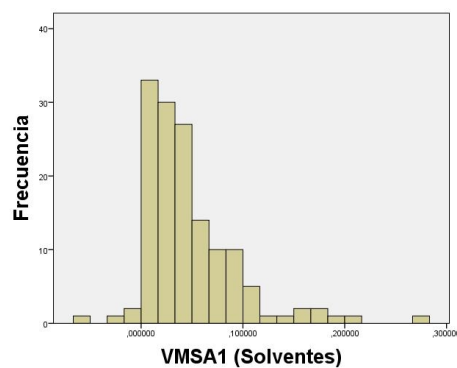
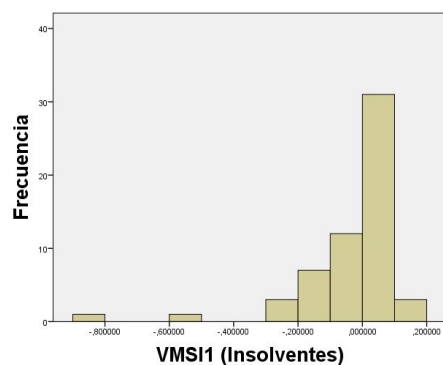
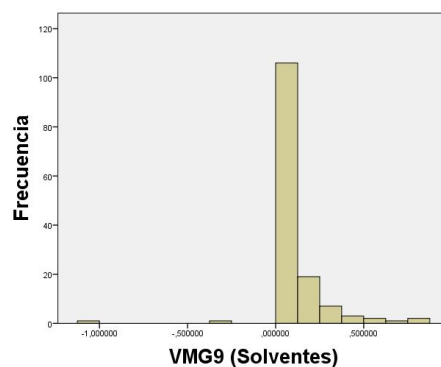
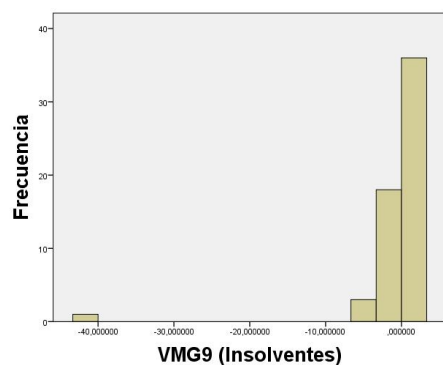
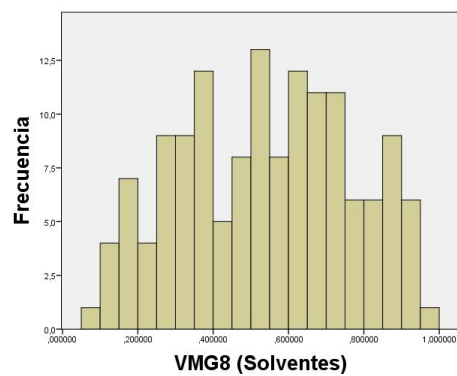
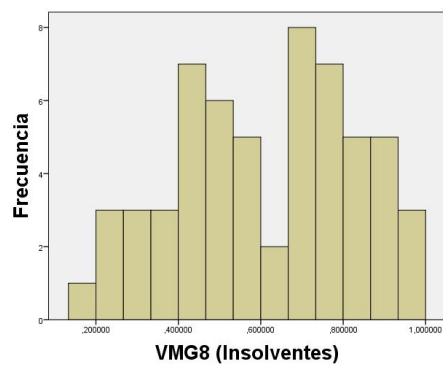


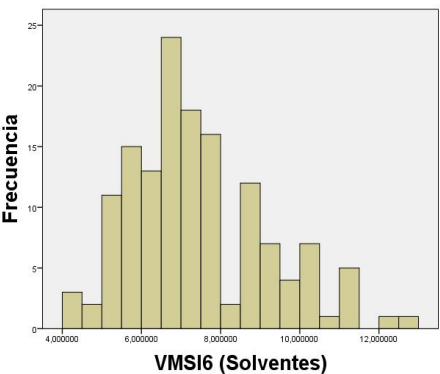
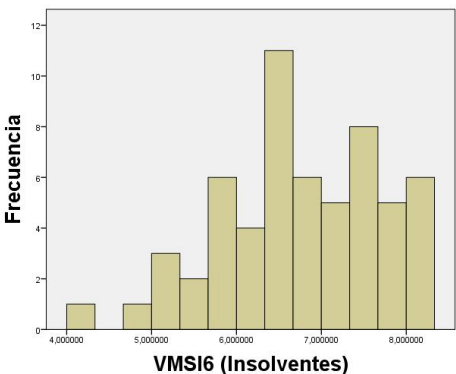
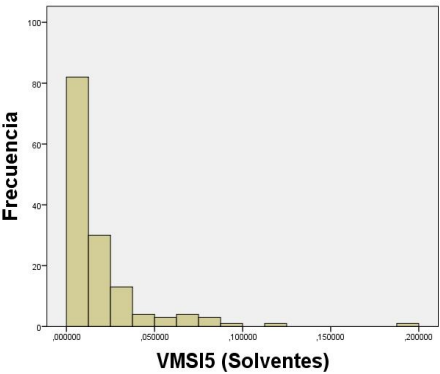
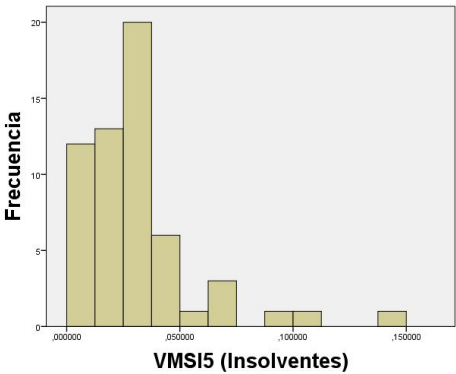
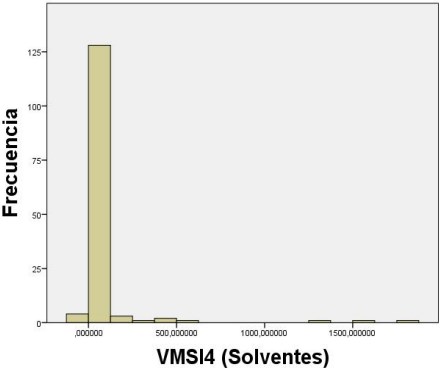
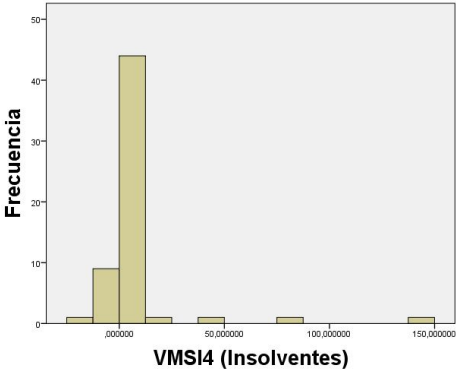
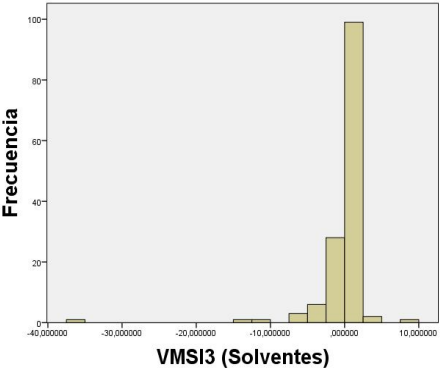
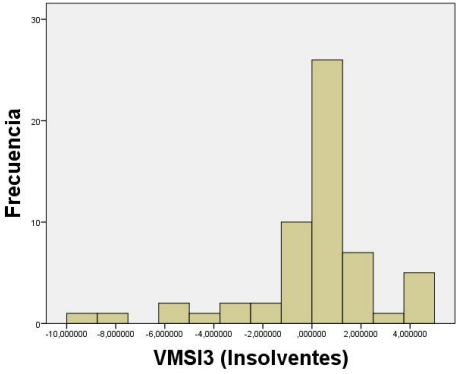


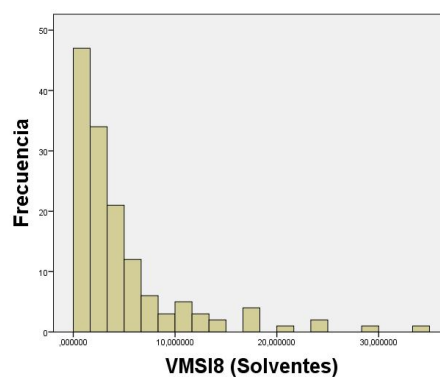
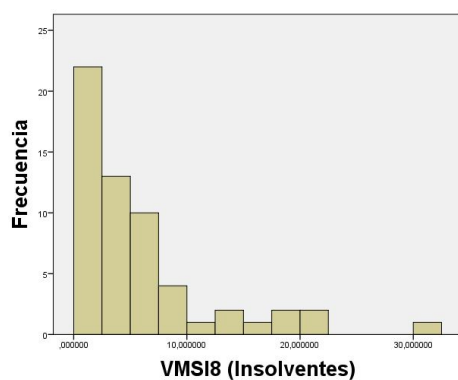
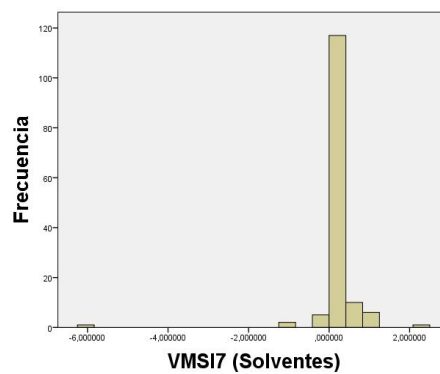
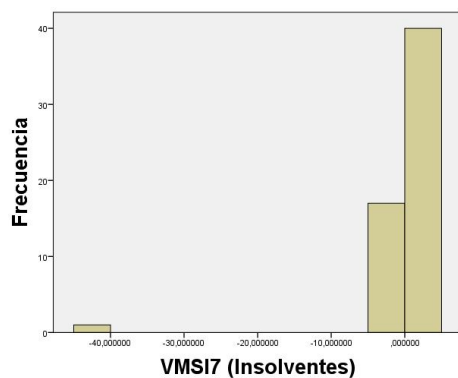
Histogramas de las variables del Sector Industria, M.II



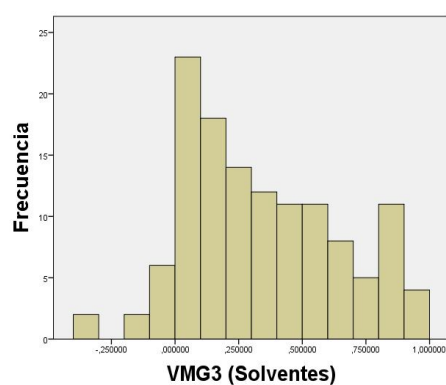
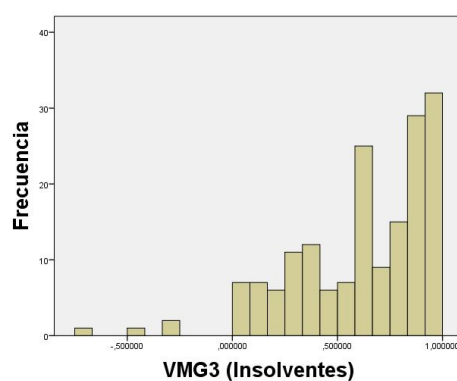
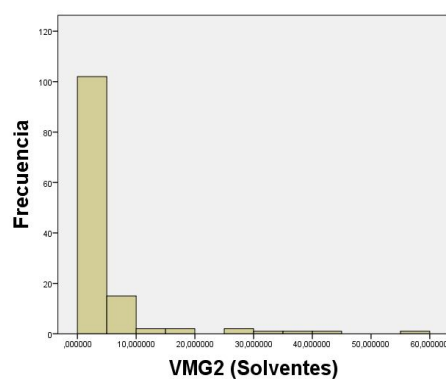
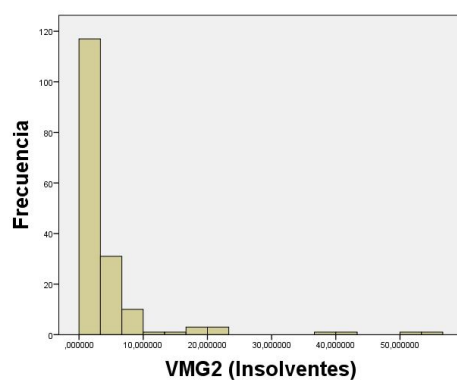
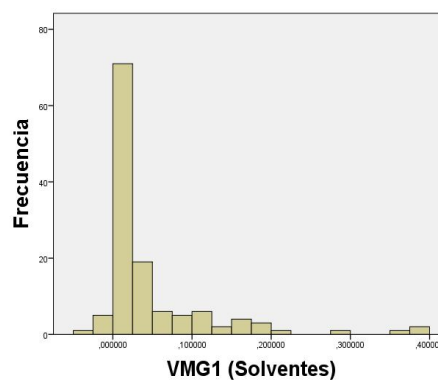
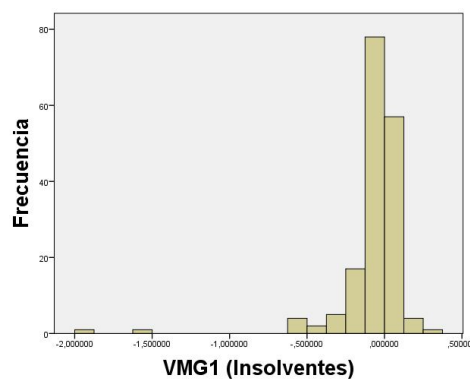


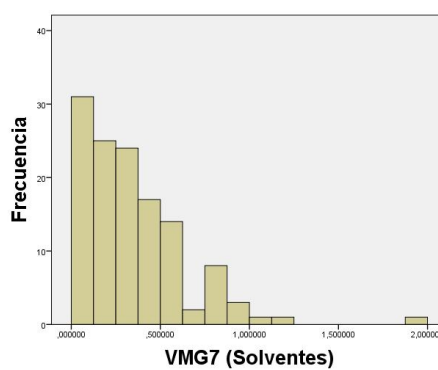
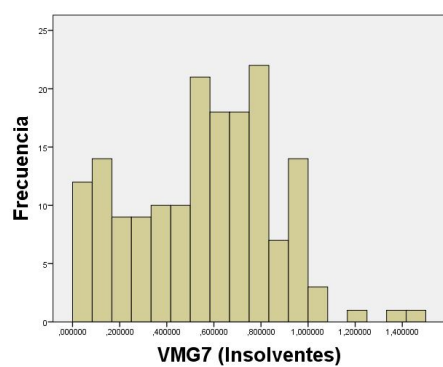
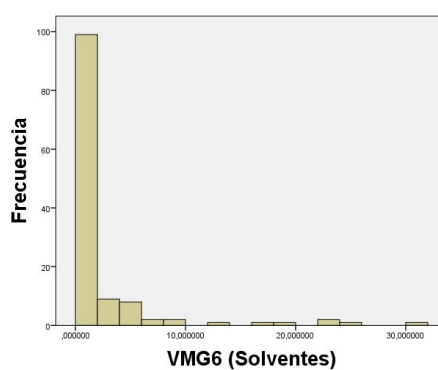
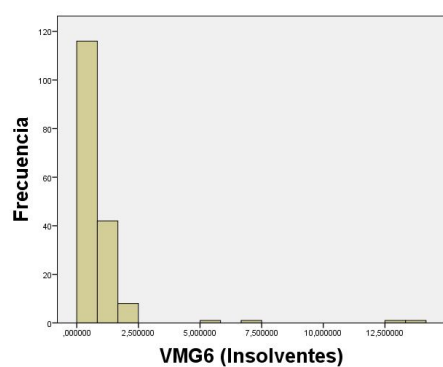
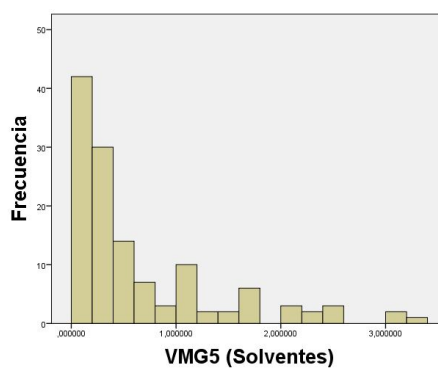
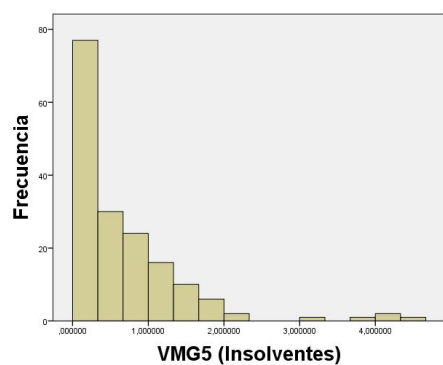
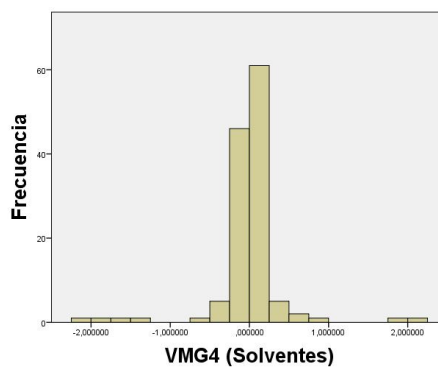
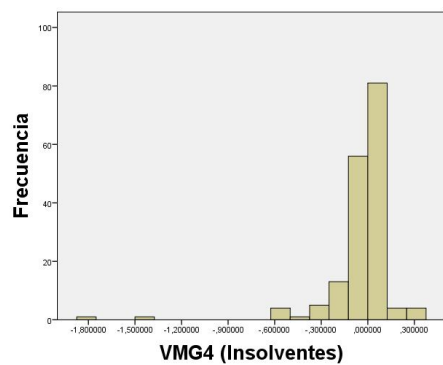


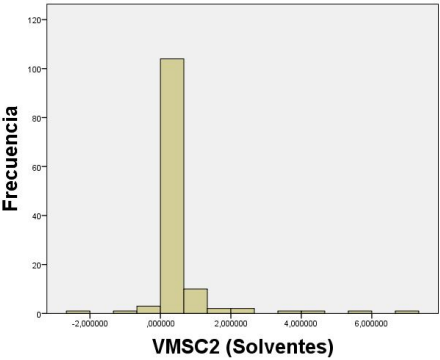
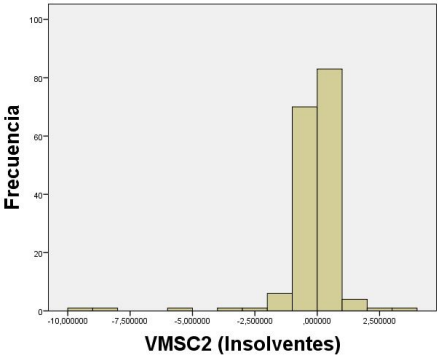
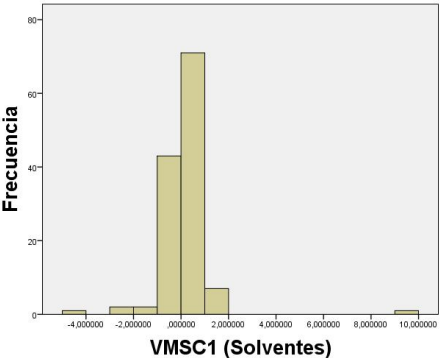
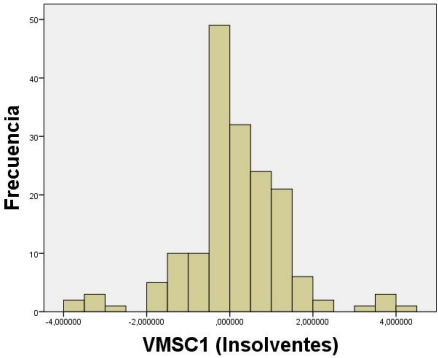
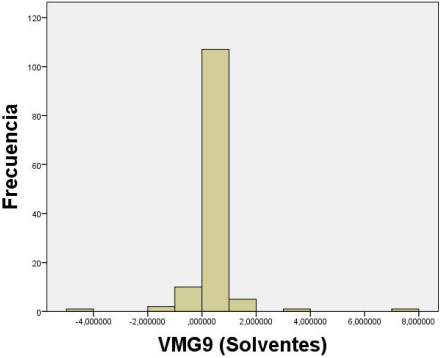
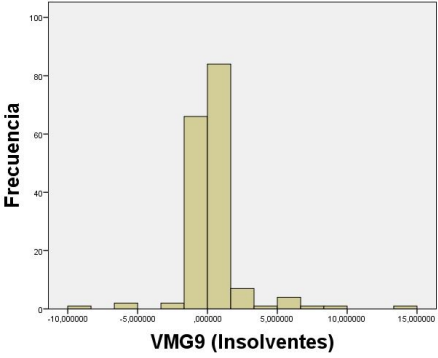
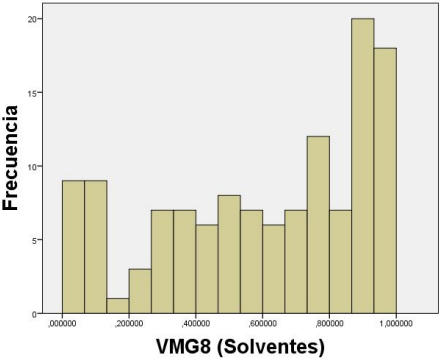
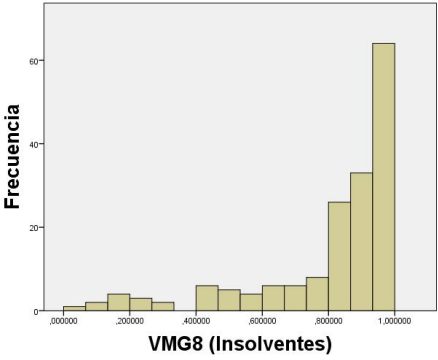


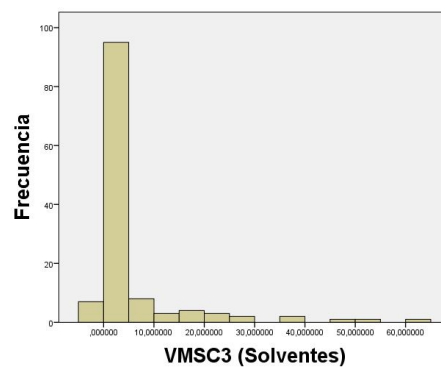
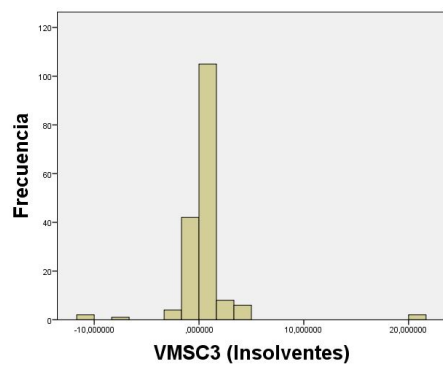


Histogramas de las variables del Sector Construcción, M.I

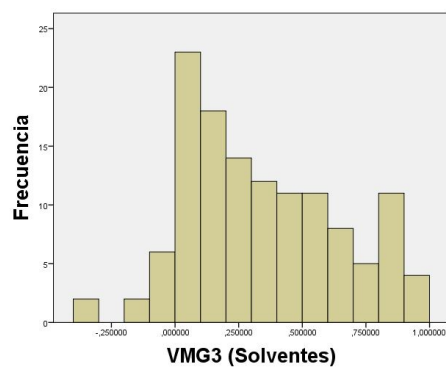
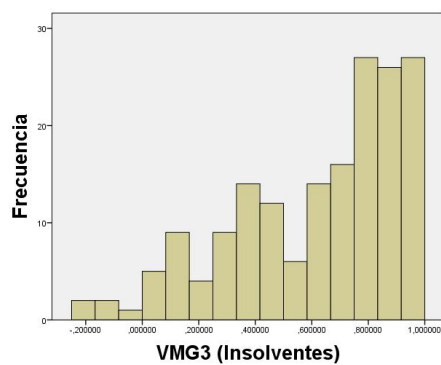
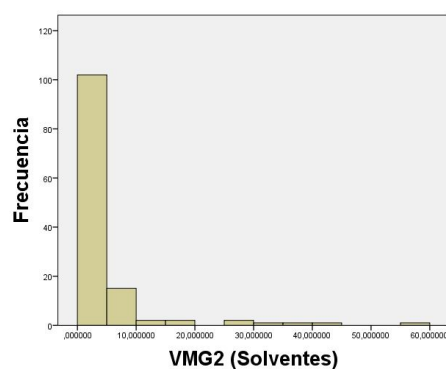
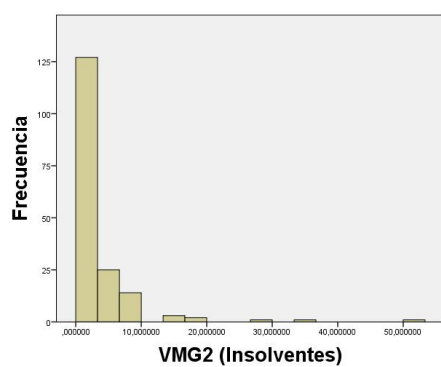
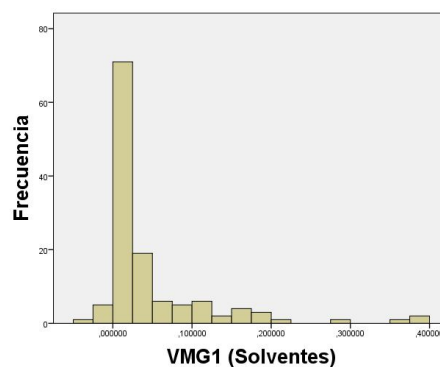
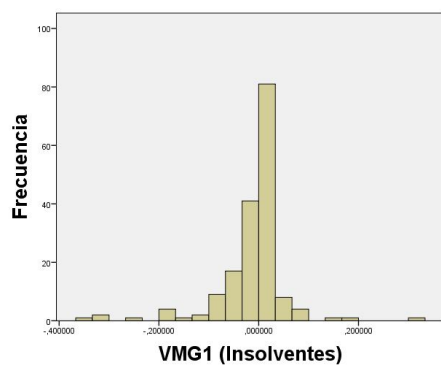


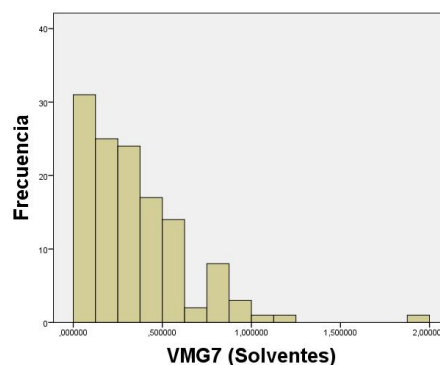
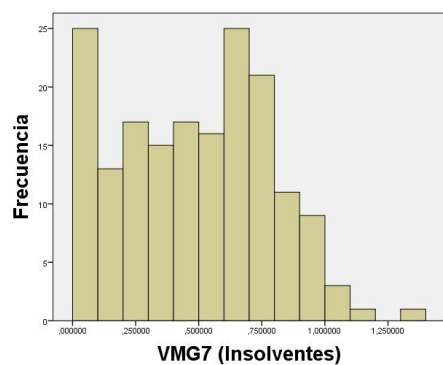
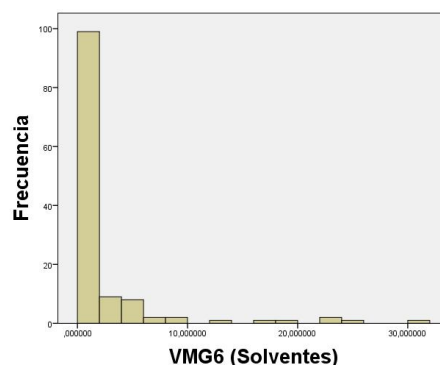
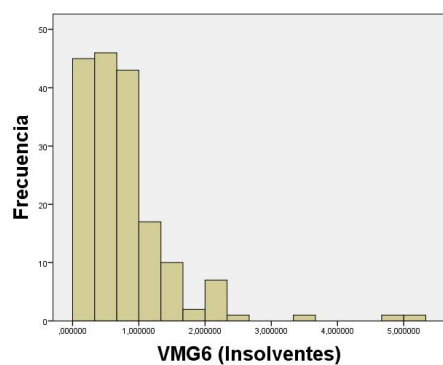
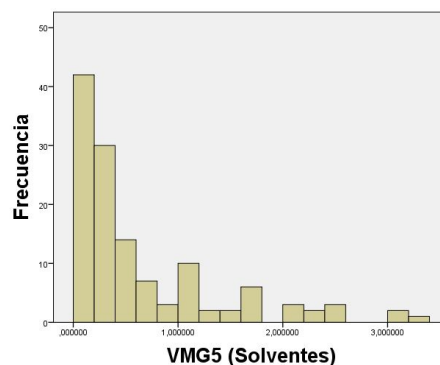
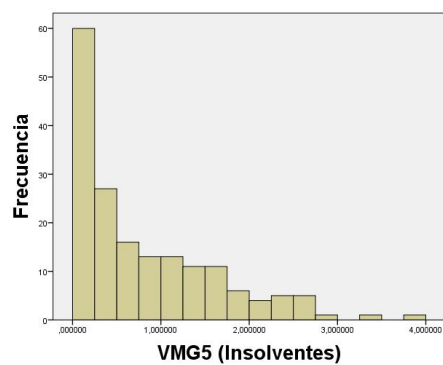
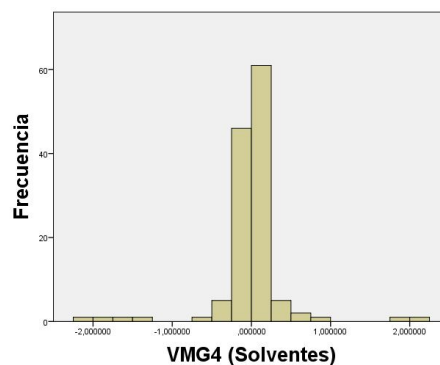
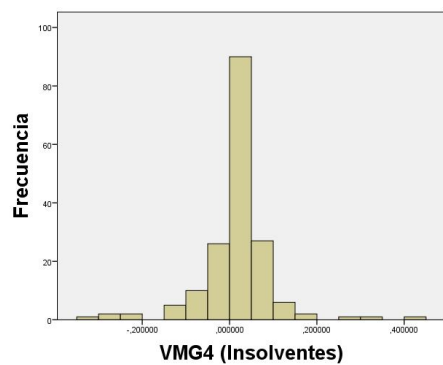


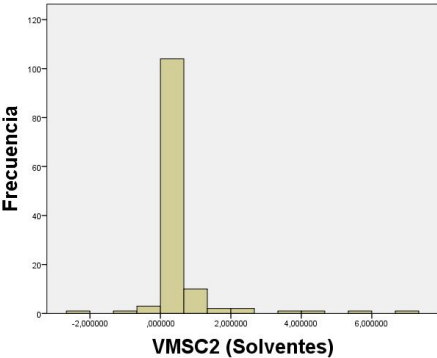
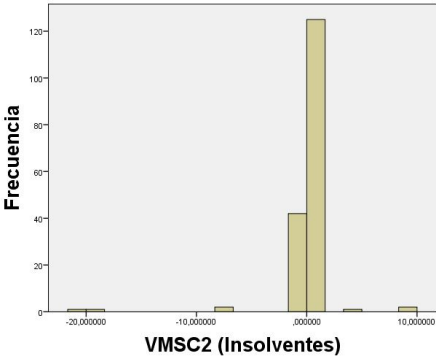
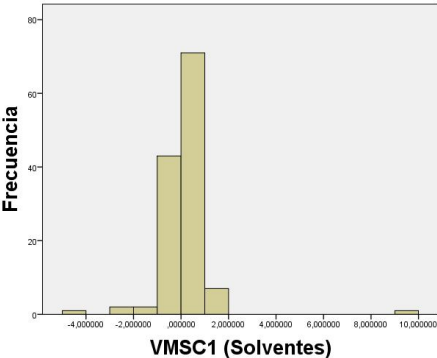
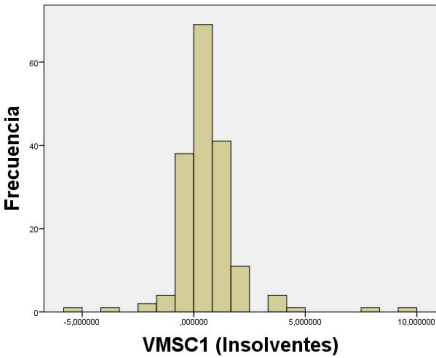
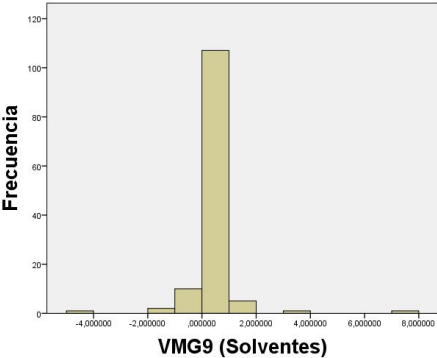
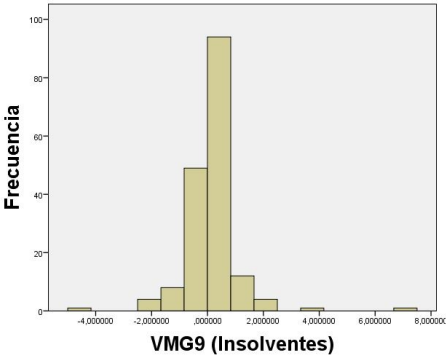
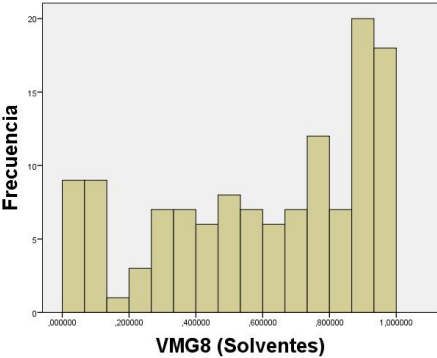
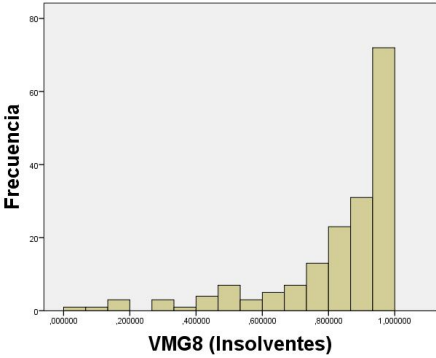


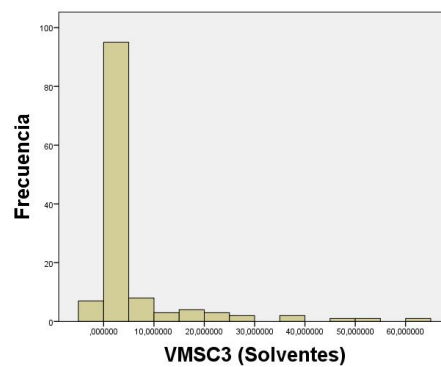
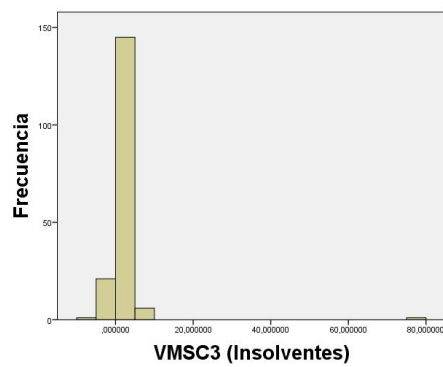


Histogramas de las variables del Sector Construcción, M.II

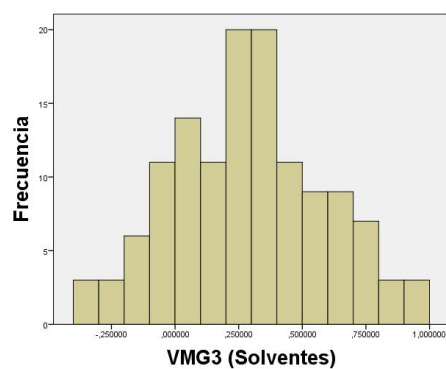
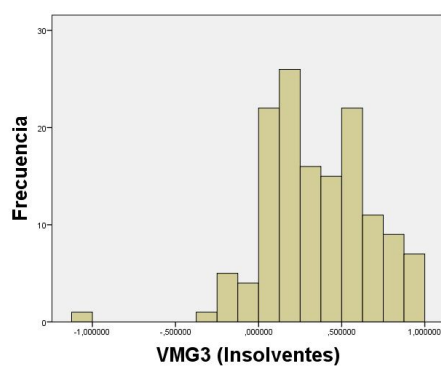
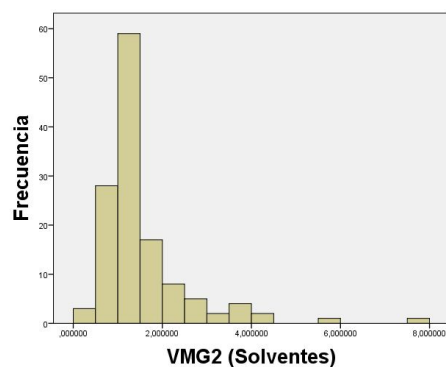
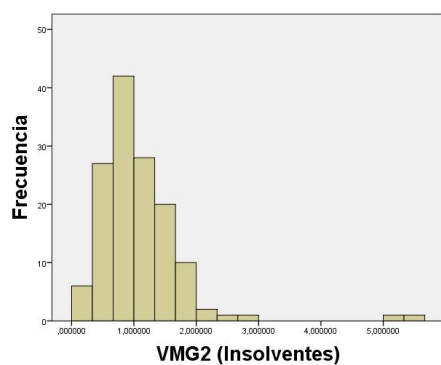
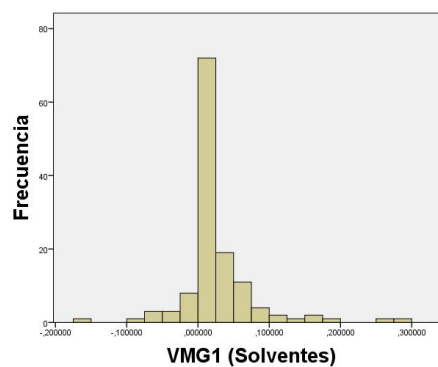
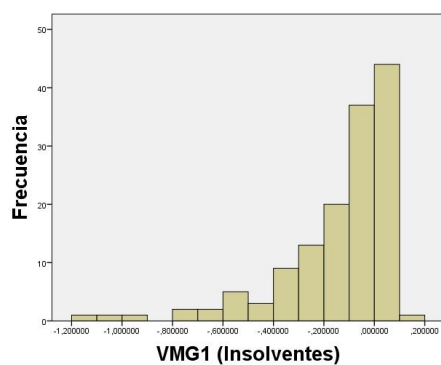


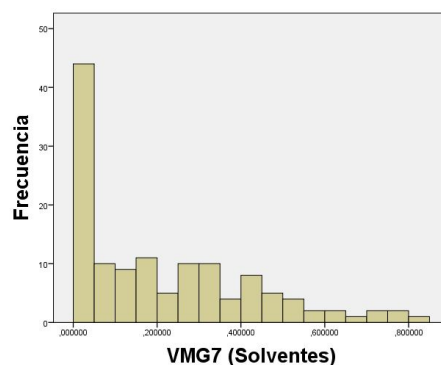
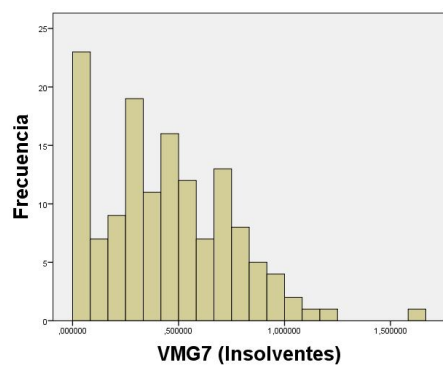
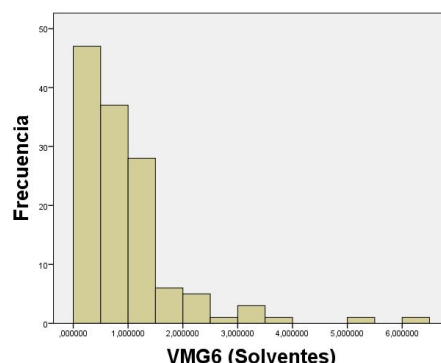
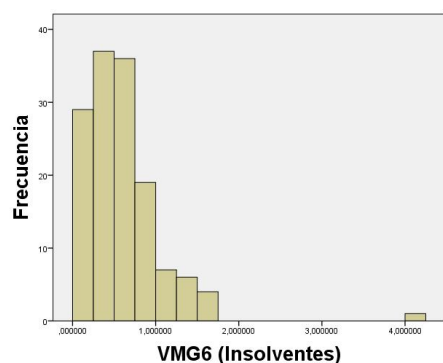
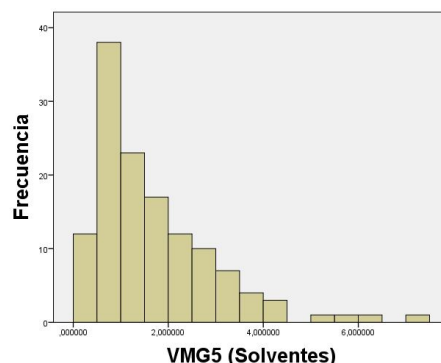
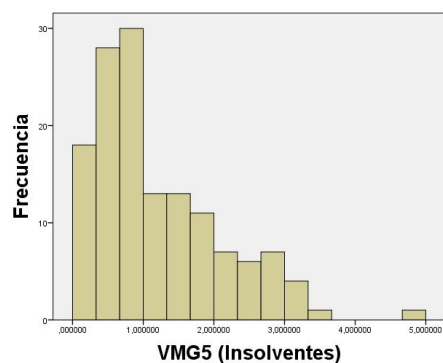
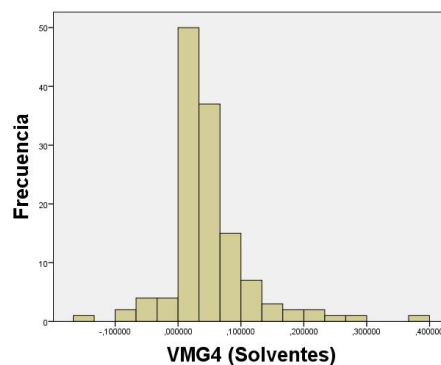
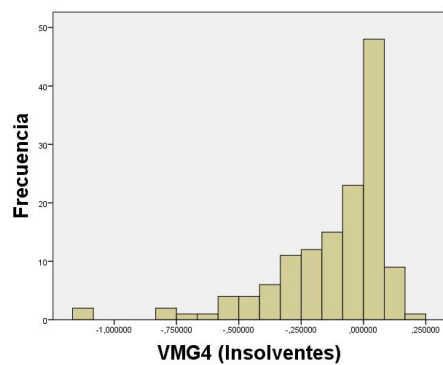


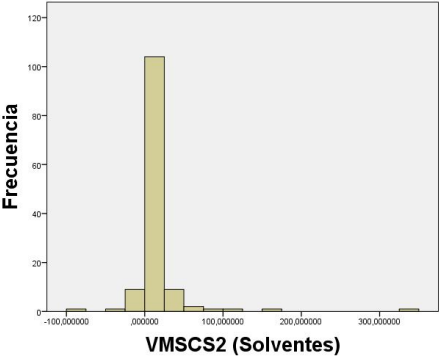
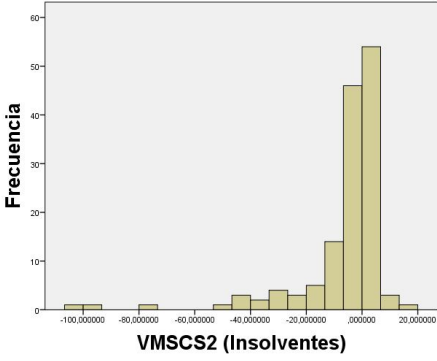
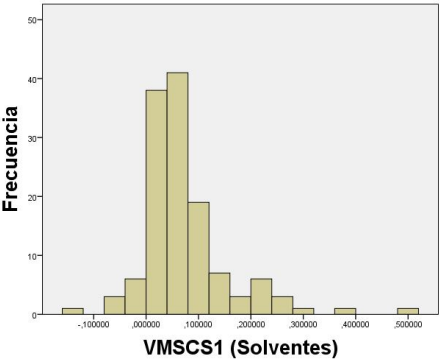
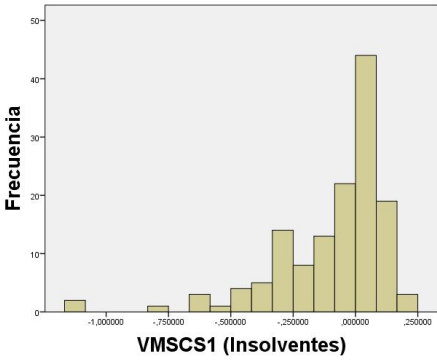
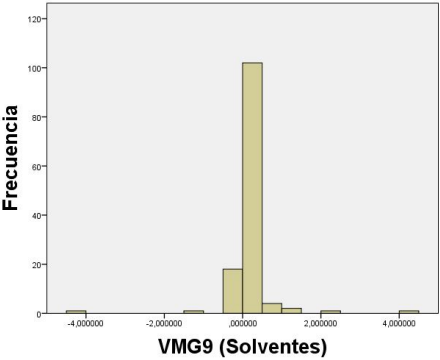
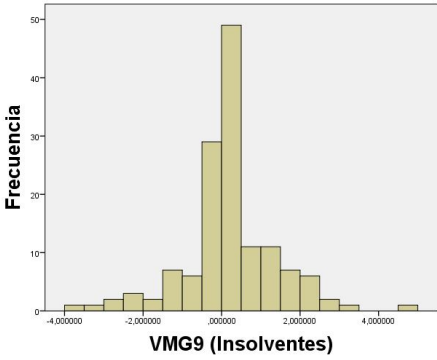
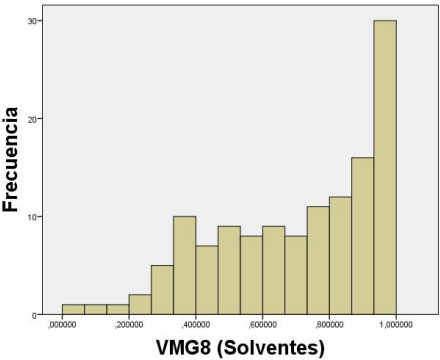
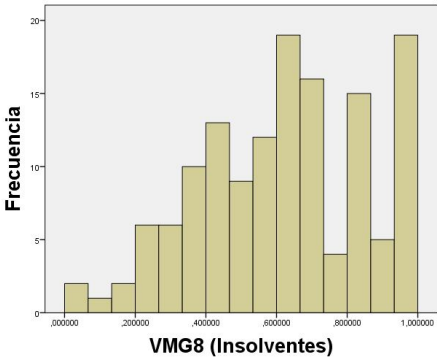


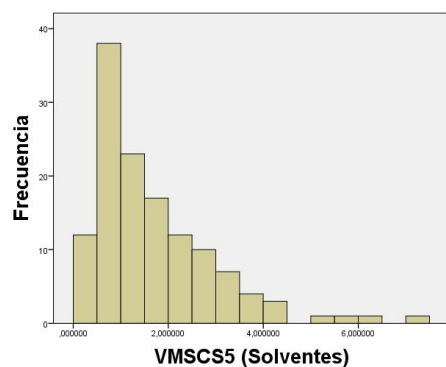
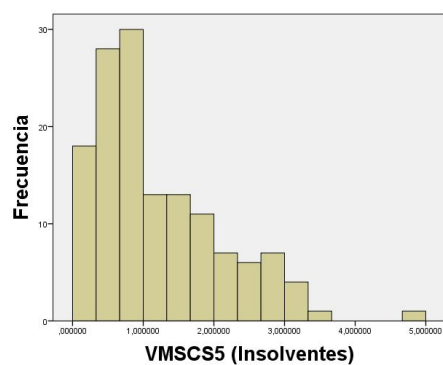
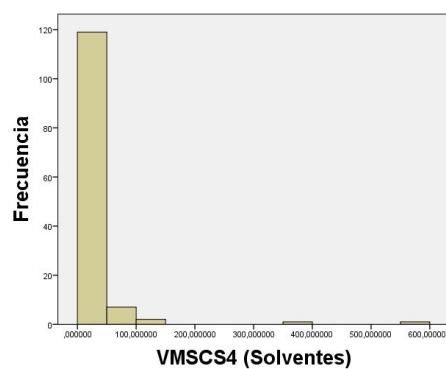
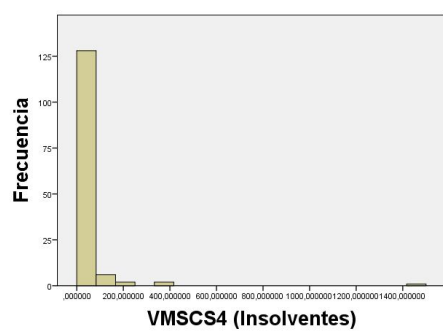
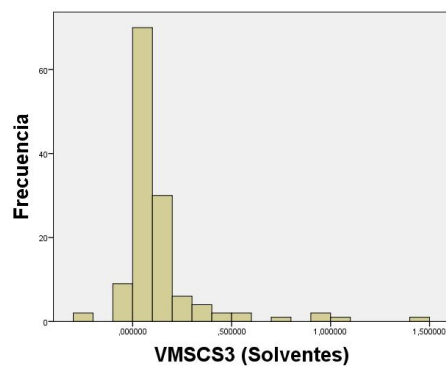
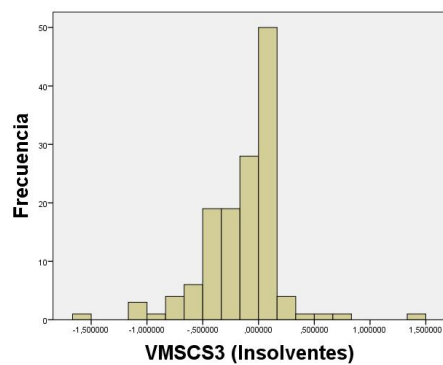


Histogramas de las variables del Sector Comercio y Servicios, M.I

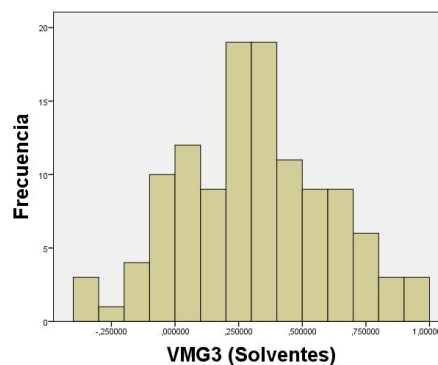
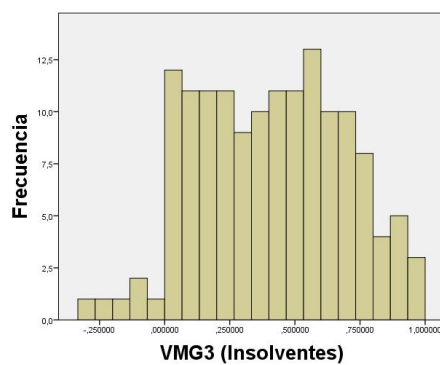
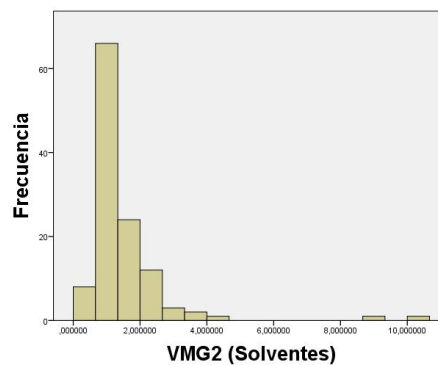
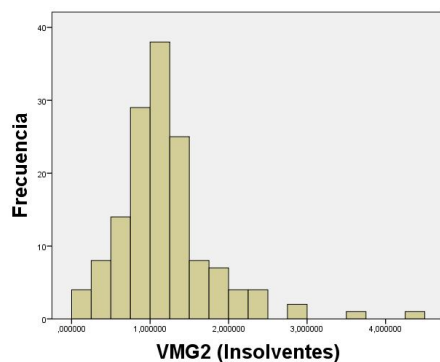
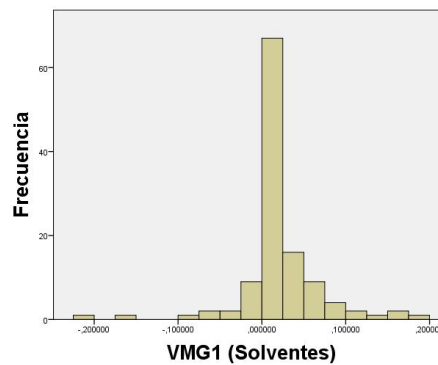
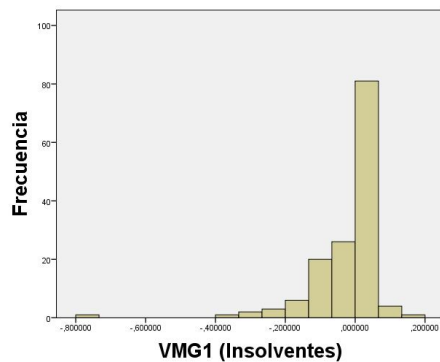


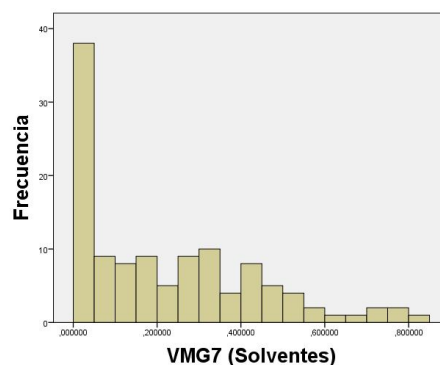
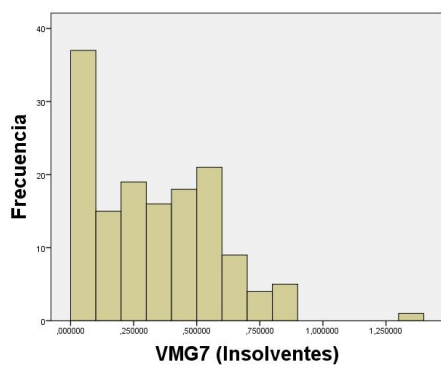
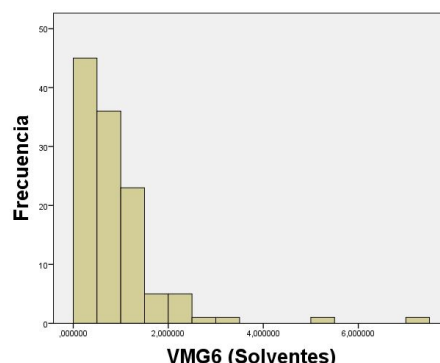
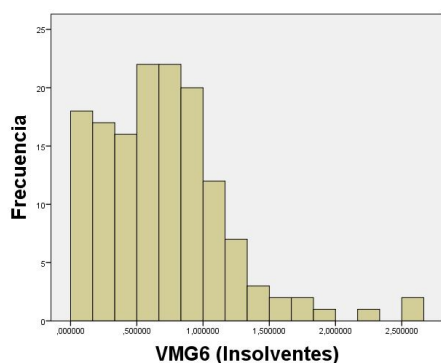
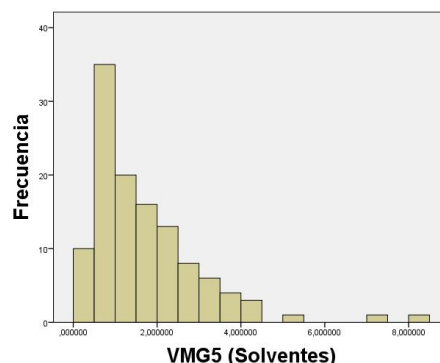
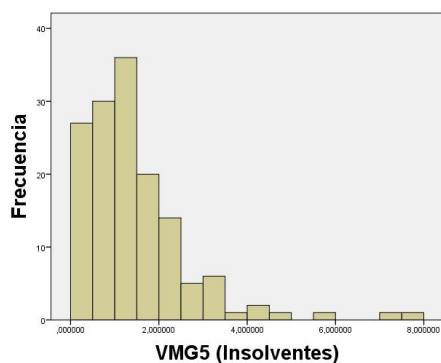
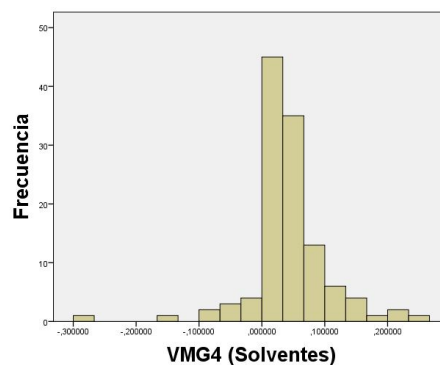
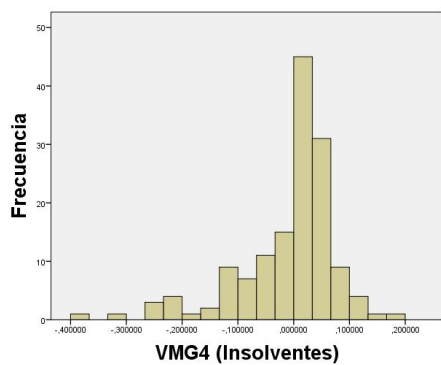


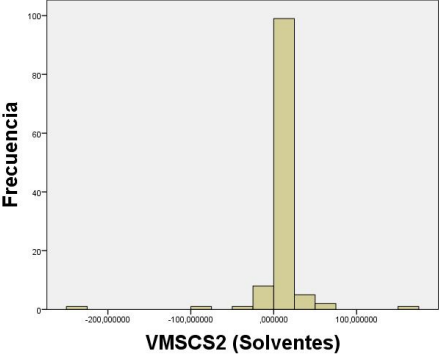
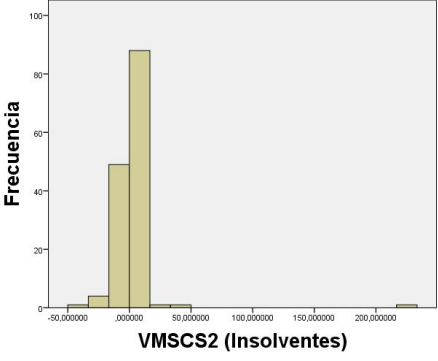
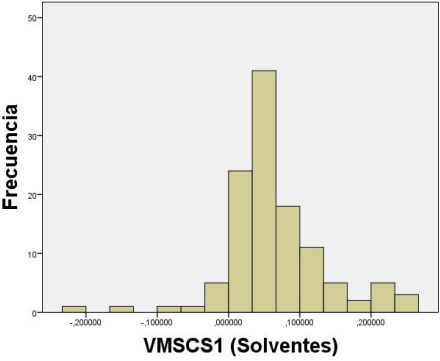
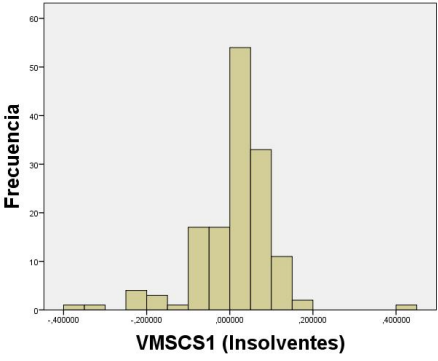
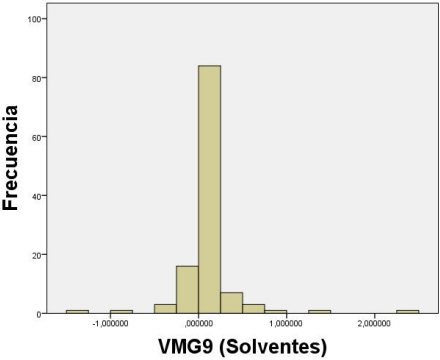
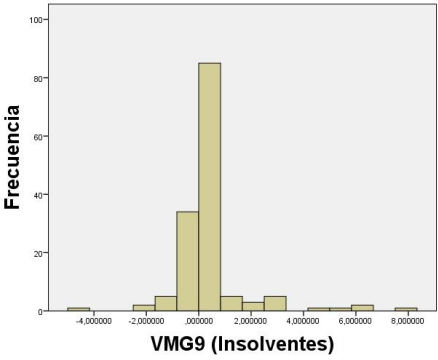
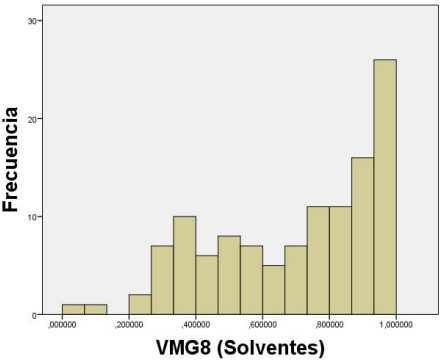
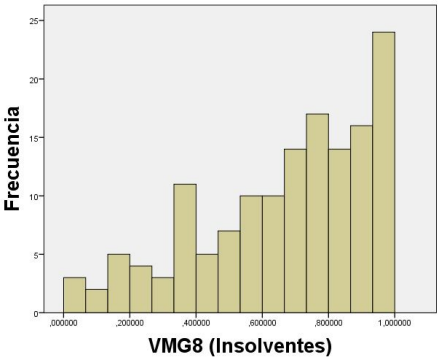


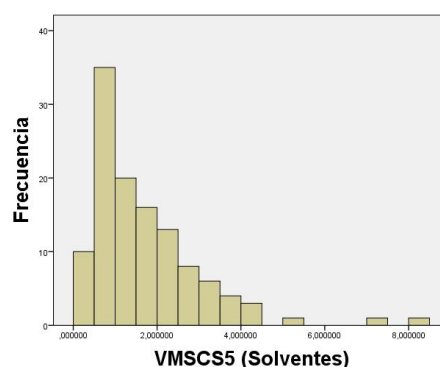
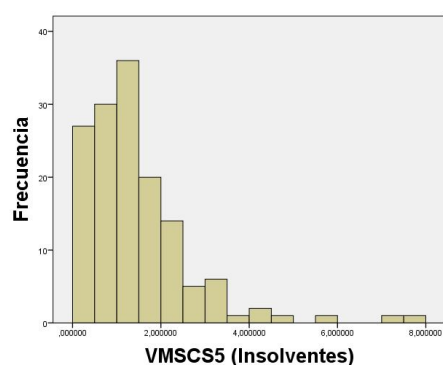
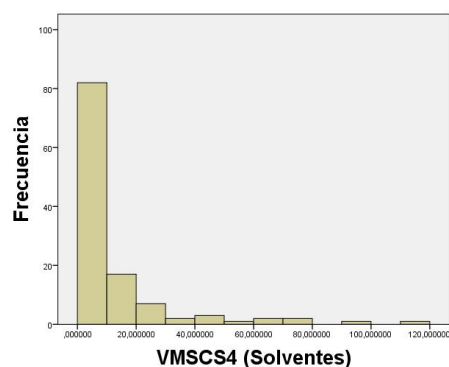
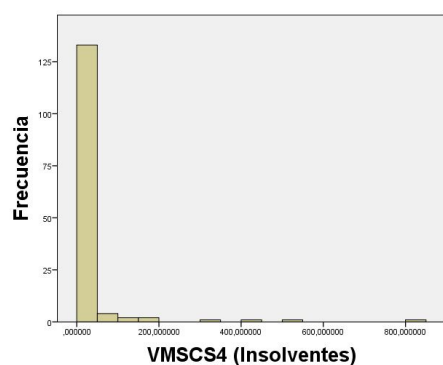
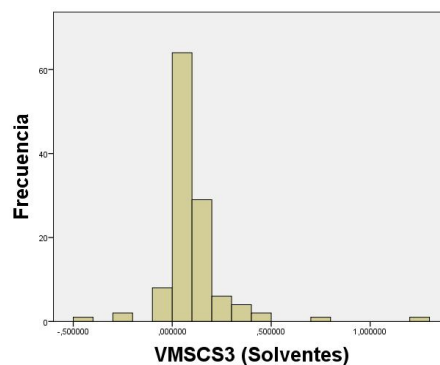
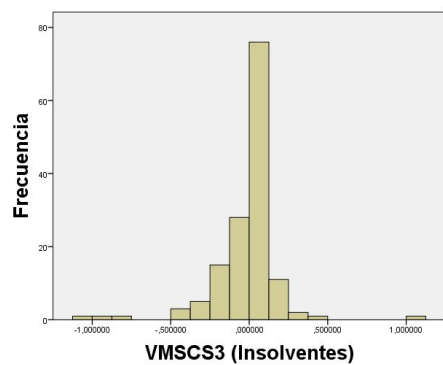


Histogramas de las variables del Sector Comercio y Servicios, M.II

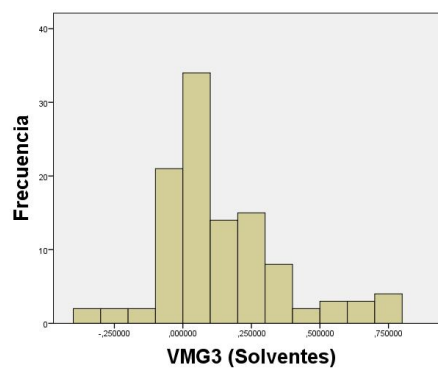
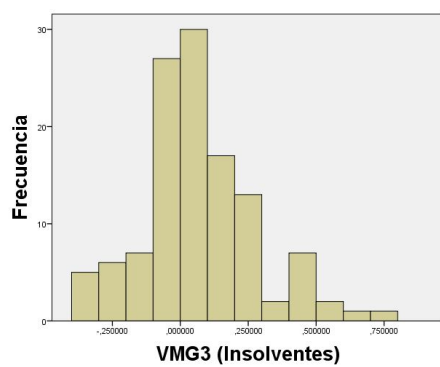
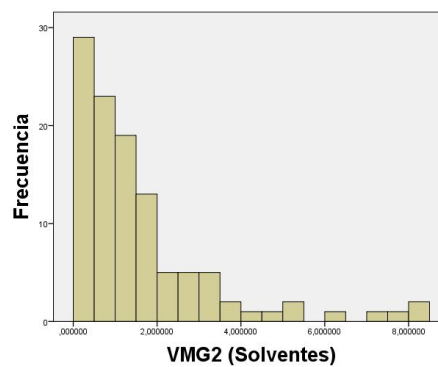
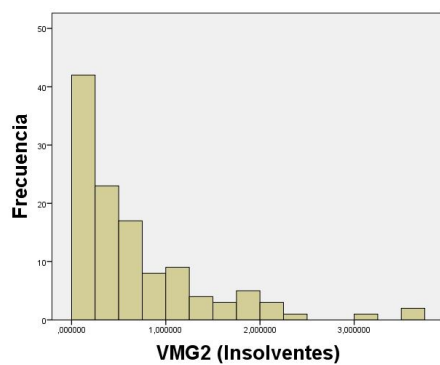
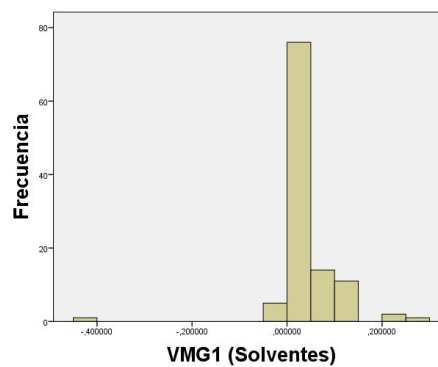
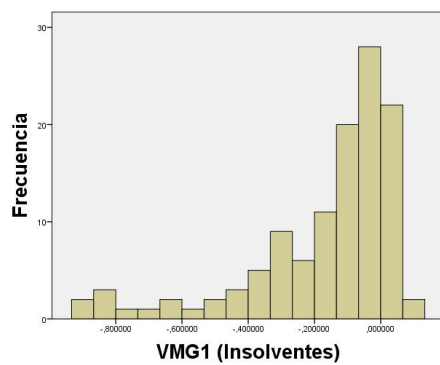


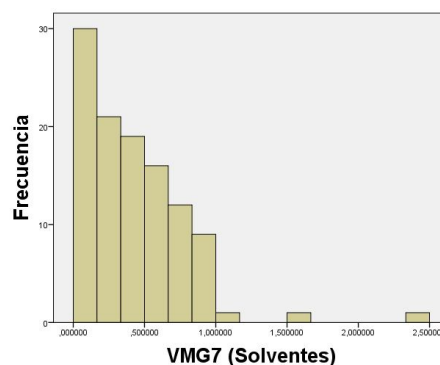
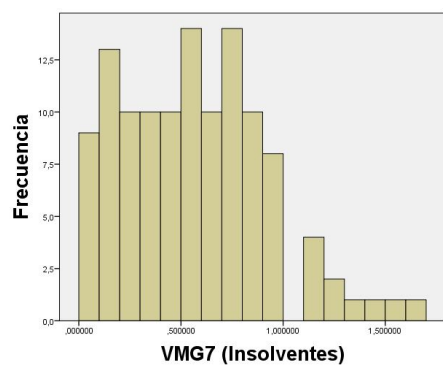
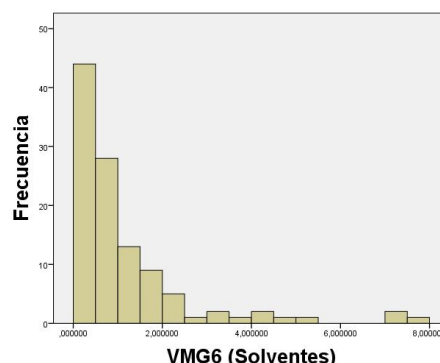
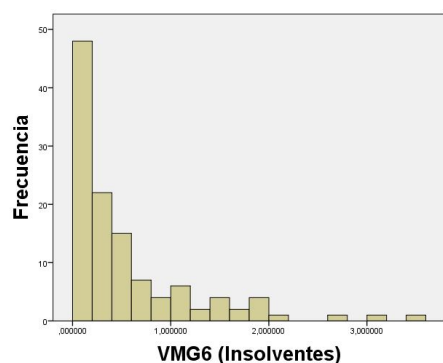
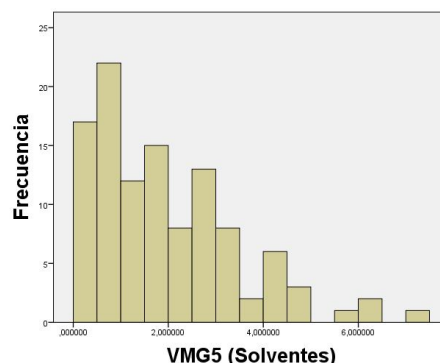
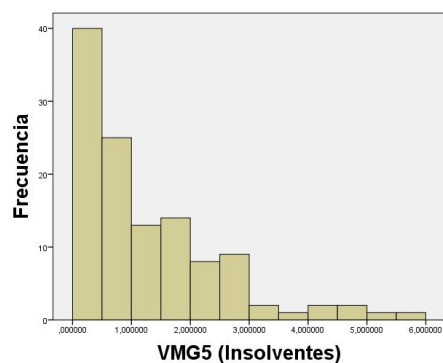
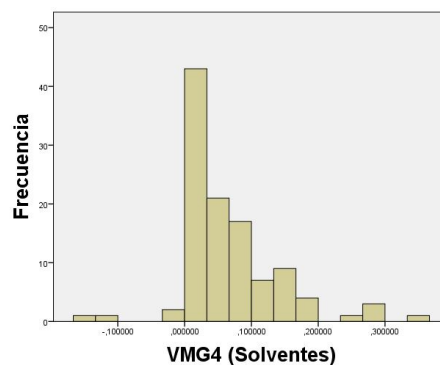
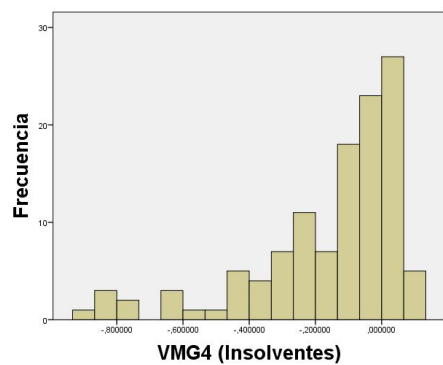


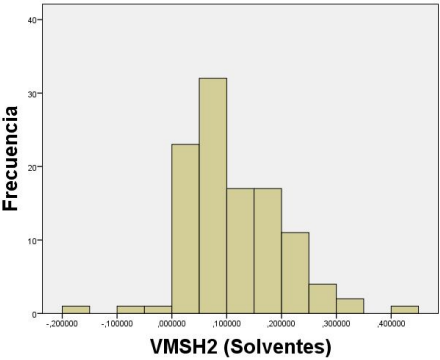
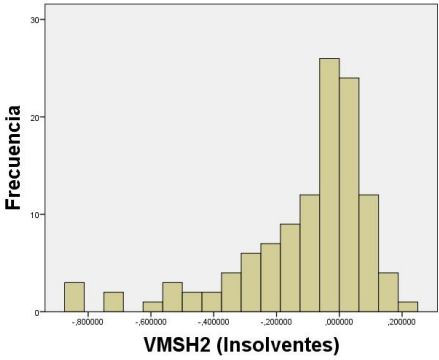
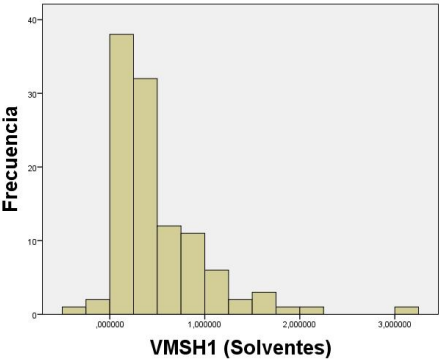
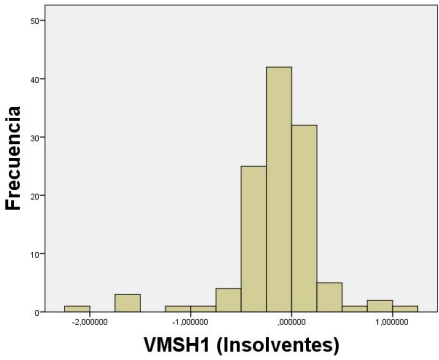
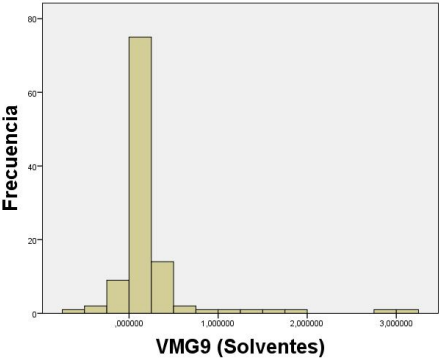
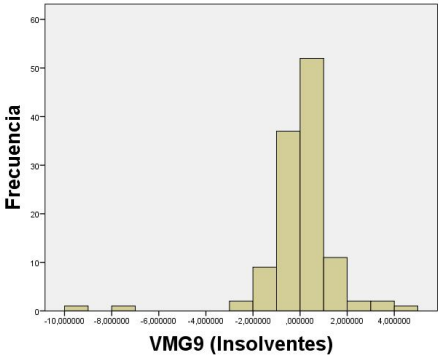
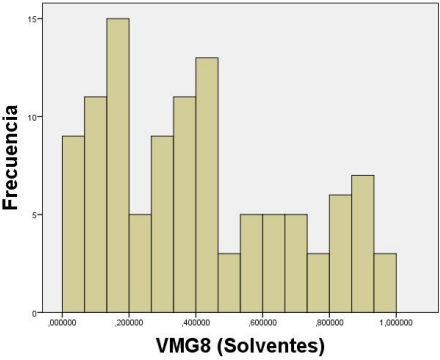
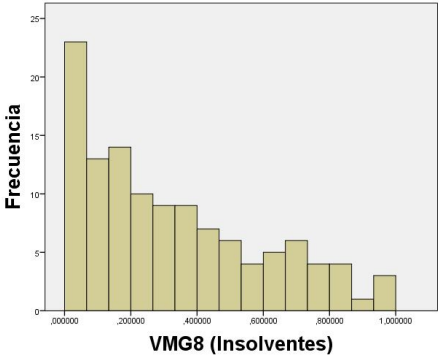


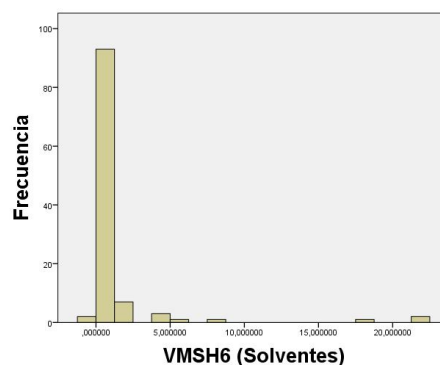
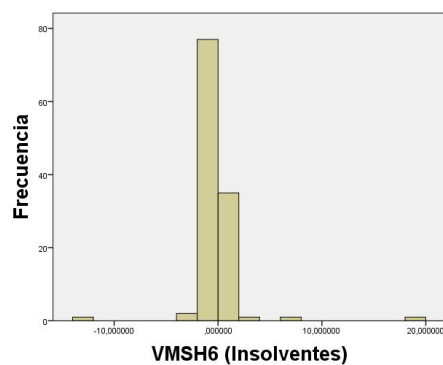
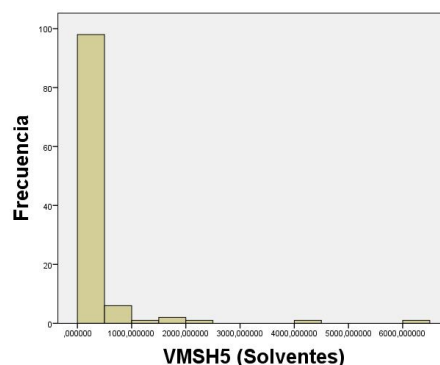
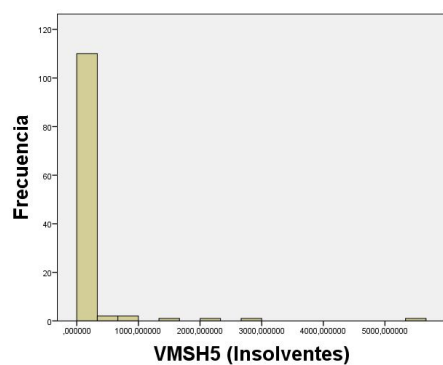
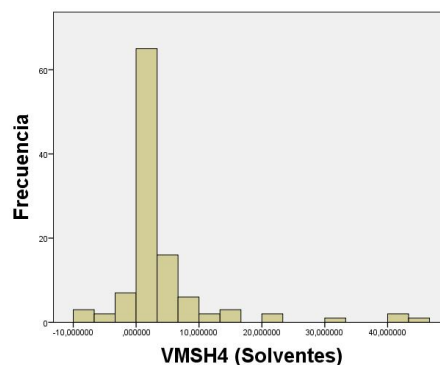
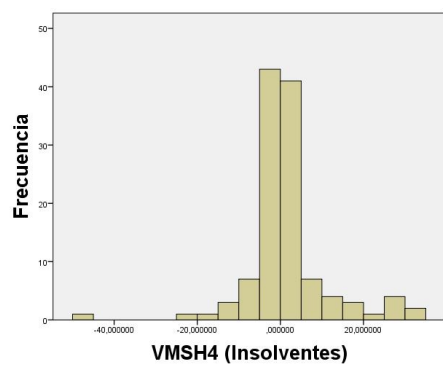
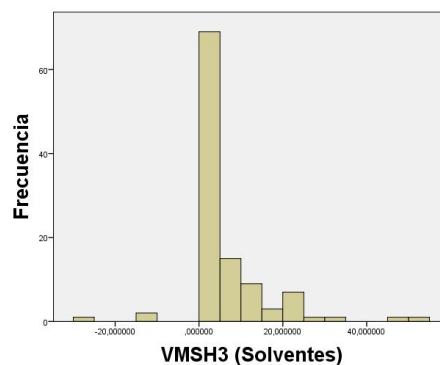
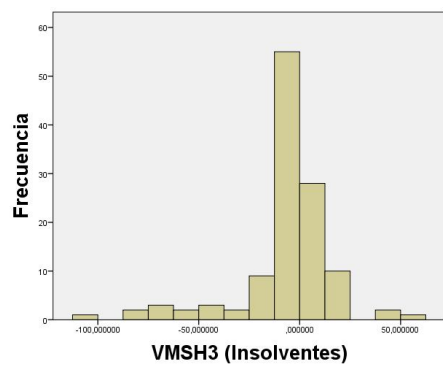


Histogramas de las variables del Sector Hostelería, M.I

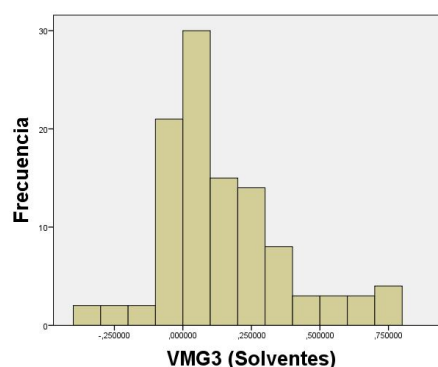
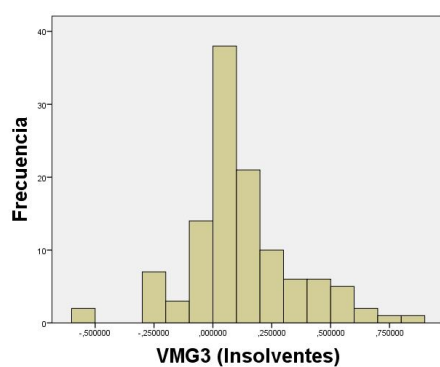
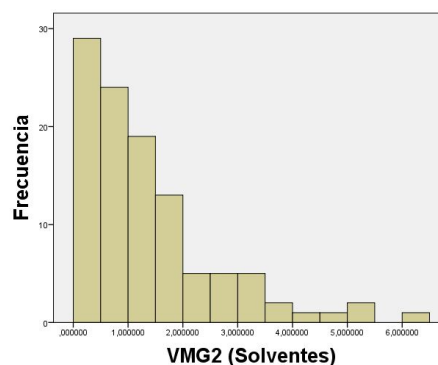
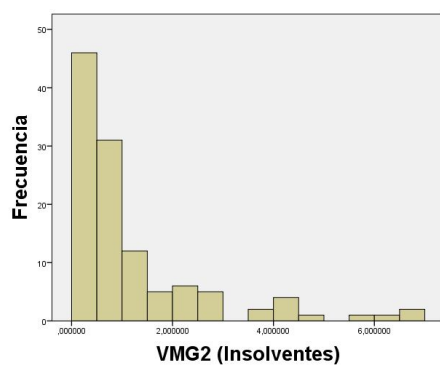
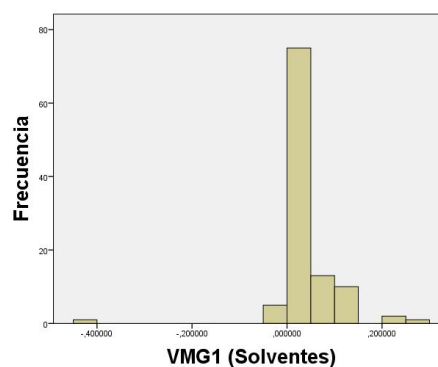
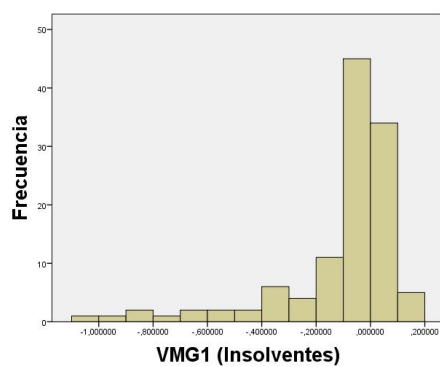


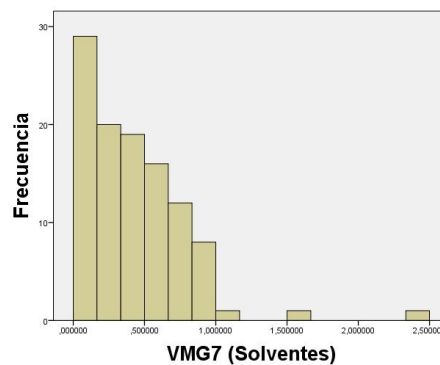
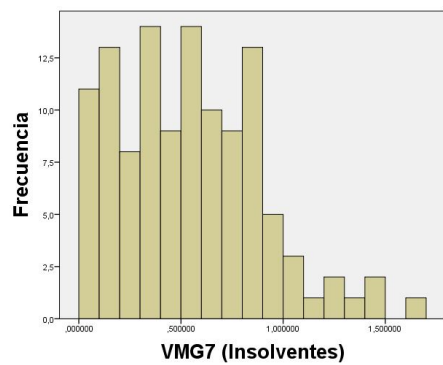
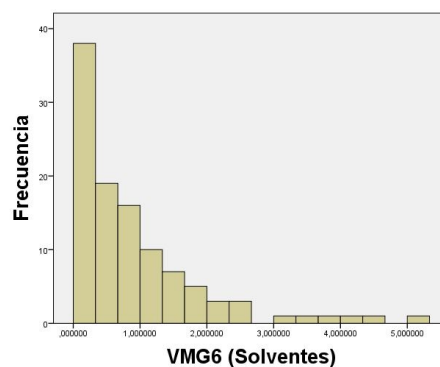
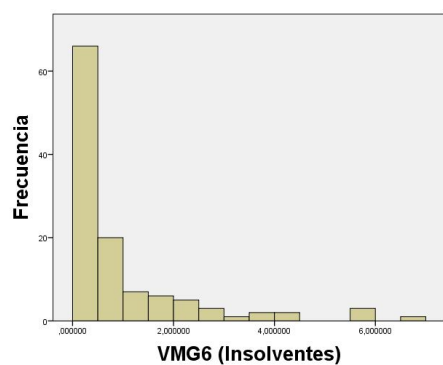
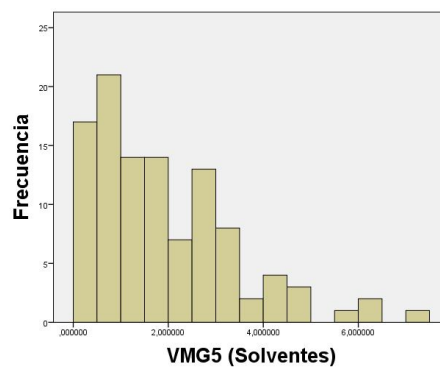
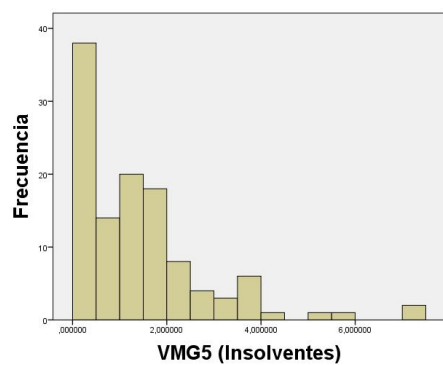
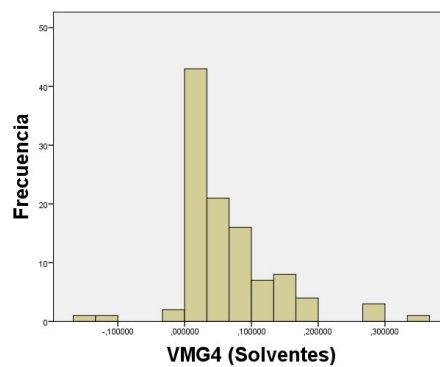
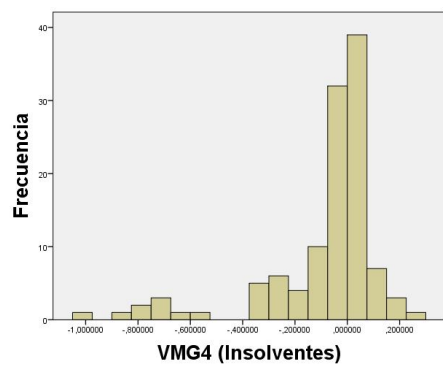


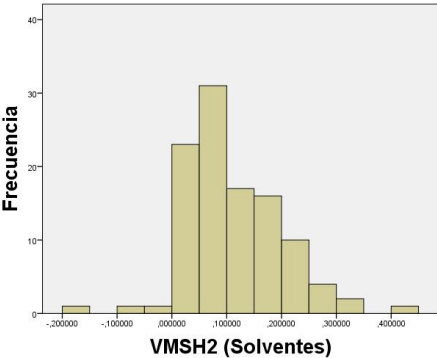
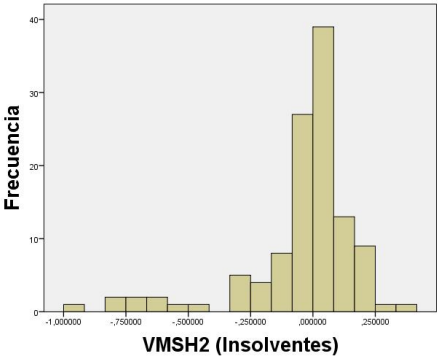
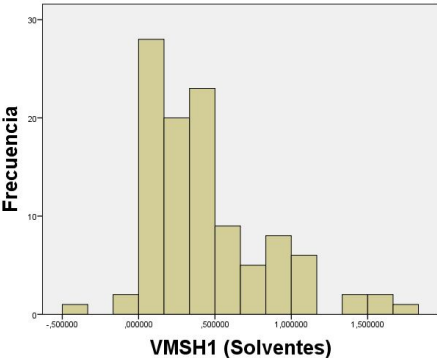
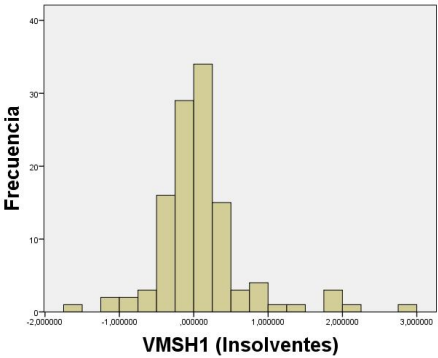
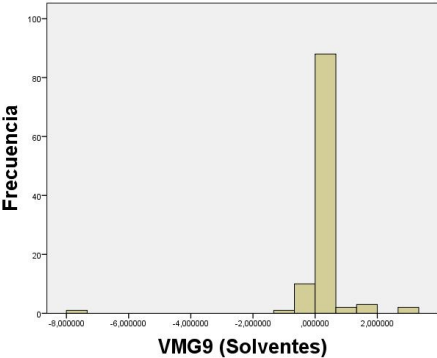
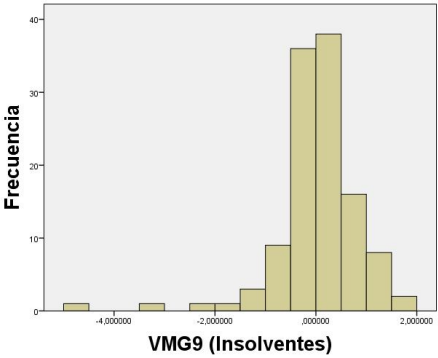
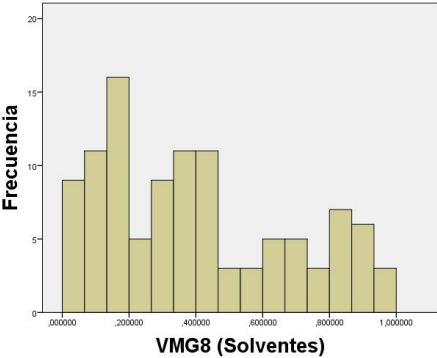
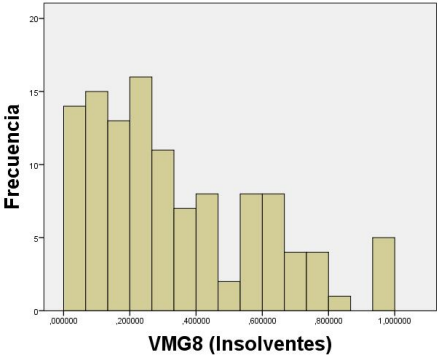


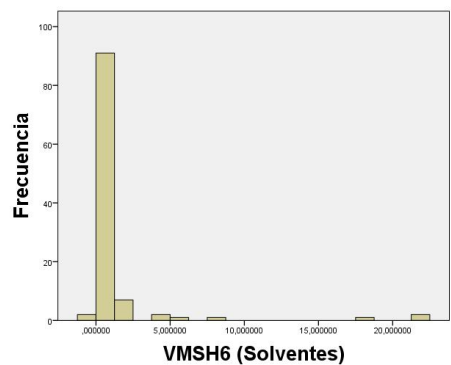
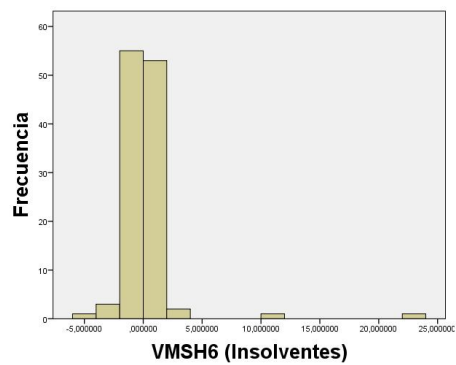
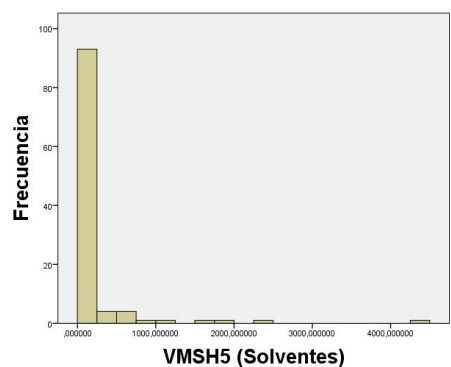
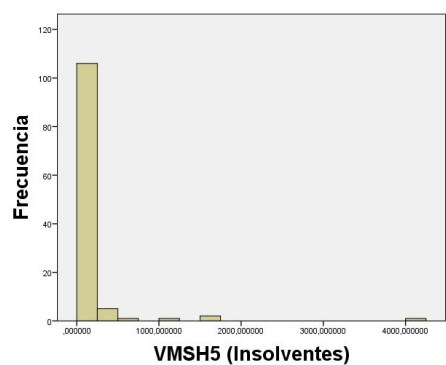
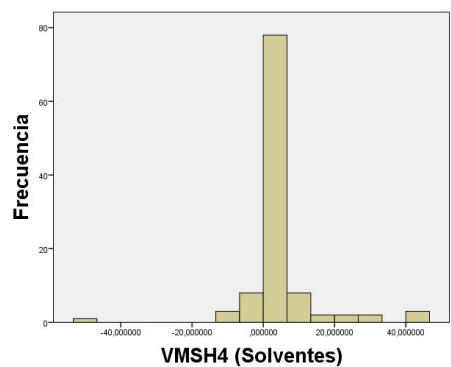
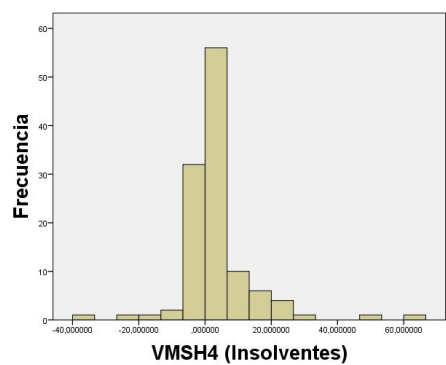
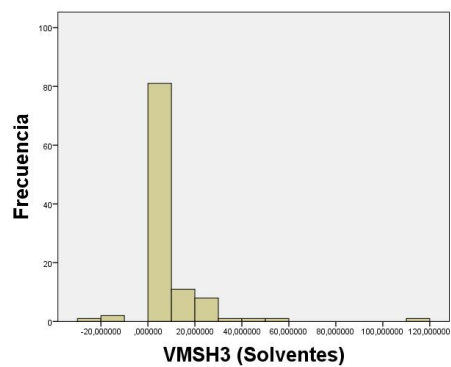
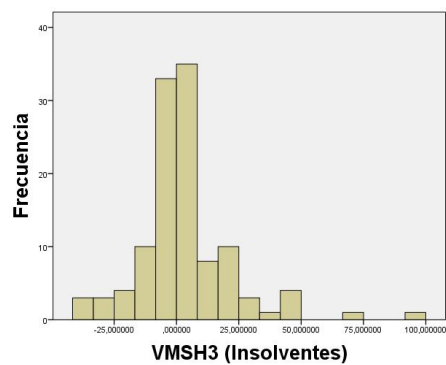


Histogramas de las variables del Sector Hostelería, M.II









ANEXO 4: GLOSARIO DE SIGLAS

MDA	Análisis Discriminante
PLS-DA	Análisis Discriminante de Mínimos Cuadrados Parciales
NN	Redes Neuronales Artificiales
EA	Algoritmos de Evolución
RS	Rough Set
CBR	Razonamiento Basado en Casos
SVM	Máquinas de Soporte Vectorial
GA	Algoritmos genéticos
GP	Programación Genética
DT	Árboles de Decisión
MCLP	Programación Lineal Multicriterio
IS	Separación Isotónica
ISO-MAP	Algoritmo de Mapeo Isométrico
CFFO	Cash Flow de las Operaciones
SIC	Standard Industrial Classification
CUSUM	Metodología de series temporales de sumas acumulativas
MLP	Perceptrón Multicapa
RFB	Red de Función Radial
BNN	Back-Propagation Neural Network
RSVM	Función radial de máquinas de soporte vectorial
DEA	Análisis Envolvente de Datos

